AMERICAN
JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. T. BELL
CALIFORNIA INSTITUTE OF TECHNOLOGY

ABRAHAM COHEN
THE JOHNS HOPKINS UNIVERSITY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

G. C. EVANS
UNIVERSITY OF CALIFORNIA

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY
HARRY BATEMAN
W. A. MANNING

MARSTON MORSE
J. R. KLINE
E. P. LANE
HARRY LEVY

ALONZO CHURCH
L. R. FORD
OSCAR ZARISKI

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LVII

1935

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

Math. Lib. C-27

MAY 6 1935

AMERICAN

JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. T. BELL
CALIFORNIA INSTITUTE OF TECHNOLOGY

ABRAHAM COHEN
THE JOHNS HOPKINS UNIVERSITY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

G. C. EVANS
UNIVERSITY OF CALIFORNIA

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY
HARRY BATEMAN
W. A. MANNING

MARSTON MORSE
J. R. KLINE
E. P. LANE
HARRY LEVY

ALONZO CHURCH
L. R. FORD
OSCAR ZARISKI

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LVII, Number 1

JANUARY, 1935

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

Further theta expansions useful in arithmetic. By M. A. Basoco and E. T. Bell, ... 1

Concerning continua of finite degree and local separating points. By G. T. Whyburn, ... 11

The topology of (path) surfaces. By Charles B. Morrey, Jr., . 17

Concerning spherical spaces. By Leonard M. Blumenthal, . . 51

On certain types of continuous transformations of metric spaces. By W. A. Wilson, ... 62

Discontinuous solutions in the non-parametric problem of Mayer in the calculus of variations. By William T. Reid, ... 69

On the inversion formula for Fourier-Stieltjes transforms in more than one dimension. By E. K. Haviland, ... 94

A note on the distribution of the zeros of the zeta-function. By Aurel Wintner, ... 101

Parallelism and equidistance in Riemannian geometry. By R. M. Peters, 103

On a rational surface of order 12 in 9-space and its projections. By B. C. Wong, ... 112

Reducible exceptional curves of the first kind. By S. F. Barber and Oscar Zariski, ... 119

The topological transformations of a simple closed curve into itself. By E. R. van Kampen, ... 142

A theory of positive integers in formal logic. Part I. By S. C. Kleene, 153

On the cooling of the earth. By Arnold N. Lowan, ... 174

The American Journal of Mathematics will appear four times yearly.

The subscription price of the Journal for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the Journal remain on sale.

Papers intended for publication in the Journal may be sent to any of the Editors.

Editorial communications may be sent to Dr. A. Cohen at The Johns Hopkins University.

Subscriptions to the Journal and all business communications should be sent to The Johns Hopkins Press, Baltimore, Maryland, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

Math.
Direct
5-7-35

FURTHER THETA EXPANSIONS USEFUL IN ARITHMETIC.

By M. A. Basoco and E. T. Bell.

1. In a paper with a similar title † several such expansions for functions of one variable are given, all the expansions in the paper being deduced from those of the so-called doubly periodic functions of the second kind,

$$\vartheta'_1 \frac{\vartheta_\alpha(x+y)}{\vartheta_\beta(x)\vartheta_\gamma(y)},$$

where the triple index $\alpha\beta\gamma$ has the values,

001,	010,	023,	032,
100,	111,	122,	133,
203,	212,	221,	230,
302,	313,	320,	331.

The complete set of these sixteen expansions (in trigonometric series) seems to have been first obtained by Hermite, although some appear implicitly in Jacobi's memoir on the rotation of a rigid body. Another method, due to Kronecker, of deriving the sixteen expansions is well known. Misprints and other inaccuracies in the final forms have been carried over to some of the standard treatises; for an exact list the reader is referred to previous papers, of which only one need be cited.‡

For the full and efficient use of the method of paraphrase (paper cited in preceding footnote), it is necessary to have the trigonometric expansions of the remaining forty-eight functions

$$K_{abc} \frac{\vartheta_a(x+y)}{\vartheta_b(x)\vartheta_c(y)},$$

where K_{abc} is a theta constant (independent of x, y), which is to be determined for a given abc other than one of the sixteen values belonging to Hermite's functions. All sixty-four expansions when used in paraphrase have the desirable effect of introducing cross-product terms into the quadratic forms appearing in the partitions. This has been amply illustrated in former papers for the sixteen quoted values of abc.

† E. T. Bell, *Messenger of Mathematics*, vol. 54 (1924), pp. 166-176.

‡ E. T. Bell, *Transactions of the American Mathematical Society*, vol. 22 (1921), pp. 1-30 and 198-219.

There remain, then, forty-eight fundamental expansions to be determined. The constant K_{abc} in a given instance, it will be noticed, is subject to arbitrary choice; a change in the constants, of course, radically alters the nature of the arithmetical functions appearing in the expansions. For choosing the 48 constants K_{abc}, there are two apparently reasonable clues—one choice being suggested by a discussion due to Krause,† the other being indicated if we require products of the type

$$K_{a\beta\gamma}\frac{\vartheta_a(x+y)}{\vartheta_\beta(x)\vartheta_\gamma(y)}\cdot K_{abc}\frac{\vartheta_a(x-y)}{\vartheta_b(x)\vartheta_c(y)}$$

to be bilinear in the squares of the twelve elliptic functions Asn, Bcn, Cns, $\cdots$ of Glaisher, where A, B, C, $\cdots$ are the classical constants $2K/\pi$, etc. The reason for the last requirement is obvious from its interpretation in terms of paraphrases. Both methods indicate the same constants. Following the suggestion of Krause, Professor D. A. F. Robinson in a forthcoming paper in the *Transactions of the Royal Society of Canada* has given the first (so far as we know) determination of these important series. If an alternative set of 48 expansions appears by the second method, we shall have an immediate source of arithmetical theorems on comparing with Robinson's. Further, either set can be used in any application of paraphrasing.

In neither set of 48 expansions has it been possible to refer only to the divisors of a simple integer, as is the case for Hermite's 16. Instead (in the present method) we need all solutions in positive integers d, δ, d', δ' of equations of the form $n = d\delta + d'\delta'$, $n = d\delta + 2d'\delta'$, etc.

Following a method suggested elsewhere ‡ we have calculated the 48 expansions very simply from the 16. This may be done in two ways, the partitions involved being, in most cases, different, so that a comparison of the two results yields, at once, a series of arithmetical results. In this paper we record the arithmetical forms of the expansions (both sets), and a sufficient indication (in § 8) of the means by which they can be recalculated. Although no applications are included, it may be mentioned that the complete set of 64 gives an endless variety of paraphrases involving functions of 2, 4, 6, 8 $\cdots$ variables, the functions being both *complete* and *incomplete*.

The 48 functions correspond to the following values of the triple index abc:

---

† Krause, *Mathematische Annalen*, vol. 30 (1887), pp. 425-436 and pp. 516-534.

‡ M. A. Basoco, *American Journal of Mathematics*, vol. 54 (April, 1932), pp. 242-252 (p. 242, footnote).

110,	213,	011,	312,
120,	223,	021,	322,
130,	233,	031,	332,
000,	303,	202,	101*,
020,	323*,	222,	121,
030,	333,	232*,	131,
300,	003*,	102*,	201,
310,	013*,	112*,	211,
330*,	033,	132,	231*,
103*,	200,	301*,	002*,
113*,	210*,	311,	012*,
123*,	220*,	321*,	022.

A star (*) indicates that the function with the starred index can be obtained by interchanging x, y in some function already in the table, and hence its expansion need not be separately calculated (except as a check). The reason for the particular grouping into sets of 12 above will appear if the expansions are verified by the method suggested in § 8. All reductions will be omitted and only the final results stated.

2. The m, n, d, δ, t, τ, notation for integers > 0 of former papers is used. Letters with suffixes or accents denote integers of the same respective kinds as those without suffixes or accents; m, τ are odd; n, d, δ, t are unrestricted (odd or even). The first Σ in a given expansion refers to *all* m or n (as defined above) in the exponent of q. The second Σ occurs in the coefficient of the general power of q, and refers to *all* solutions, for m or n fixed, of the indicated equations (or partitions) at the head of a particular set of expansions.

For easy reference we state all the partitions here:

(I) $m = 2n_1 + m_2;\quad m = t\tau,\quad n_1 = t_1\tau_1,\quad m_2 = t_2\tau_2.$

(II) $n = n_1 + n_2;\quad n = t\tau,\quad n_1 = t_1\tau_1,\quad n_2 = t_2\tau_2.$

(III) $\begin{cases} m = 2n_1 + n_2; & m = t\tau, \quad n_1 = d_1\delta_1, \quad n_2 = t_2\tau_2, \\ 2n = 2n'_1 + n'_2; & 2n = t'\tau', \quad n = d\delta, \quad n'_1 = d'_1\delta'_1, \quad n'_2 = t'_2\tau'_2. \end{cases}$

(IV) $m = m_1 + 4n_2;\quad m = t\tau,\quad m_1 = t_1\tau_1,\quad n_2 = d_2\delta_2.$

(V) $2n = 2n_1 + 2n_2;\quad n = d\delta,\quad n_1 = d_1\delta_1,\quad n_2 = d_2\delta_2.$

We shall write

$$\phi_{abc}(x,y) \equiv K_{abc}\psi_{abc}(x,y), \quad \psi_{abc}(x,y) \equiv \frac{\vartheta_a(x+y)}{\vartheta_b(x)\vartheta_c(y)},$$

$(-1/h) \equiv (-1)^{(h-1)/2}$ if h is an odd integer. Since

$$\phi_{abc}(x,y) = \phi_{acb}(y,x),$$

it is sufficient to state the expansion of one of these functions. The constants K_{abc} are to be specified with the expansions.

3. There are 20 expansions in which the partition is (I) of § 2. All are of the form

$$\phi_{abc}(x,y) = 4\sum_{(m)} q^{m/2}(\sum F_{abc}(x,y;m)),$$

the first Σ referring to all m (in the notation explained in § 2), the second to all integers defined by the partition (I) for m fixed; namely, $\Sigma F_{abc}(x,y;m)$ is the coefficient of $q^{m/2}$. To state the expansions concisely it is sufficient to tabulate K_{abc} and $F_{abc}(x,y;m)$ for the 20 indices abc concerned. Attending to the remark at the end of § 2, we need only the following 13.

	abc	K_{abc} (Partition (I))
(1)	110	$\vartheta'_1\vartheta_2\vartheta_3$
(2)	120	$\vartheta'_1\vartheta_0\vartheta_2$
(3)	130	$\vartheta'_1\vartheta_0\vartheta_3$
(4)	213	$\vartheta'_1\vartheta_2\vartheta_3$
(5)	223	$\vartheta'_1\vartheta_0\vartheta_2$
(6)	233	$\vartheta'_1\vartheta_0\vartheta_3$
(7)	000	$\vartheta'_1\vartheta_2\vartheta_3$
(8)	030	$\vartheta'_1\vartheta_0\vartheta_2$
(9)	303	$\vartheta'_1\vartheta_2\vartheta_3$
(10)	333	$\vartheta'_1\vartheta_0\vartheta_2$
(11)	300	$\vartheta'_1\vartheta_0\vartheta_2$
(12)	033	$\vartheta'_1\vartheta_2\vartheta_3$
(13)	200	$\vartheta'_1\vartheta_0\vartheta_3$.

The $F_{abc}(x,y;m)$ are numbered correspondingly and are as follows:

(1) $\sin(tx+\tau y)\csc x - 2[\cos\{(\tau_1+t_2)x+\tau_2 y\} - \cos\{(\tau_1-t_2)x-\tau_2 y\}]$.

(2) $\sin(tx+\tau y)\sec x$
$+ 2(-1)^{m_1}(-1/\tau_1)[\sin\{(\tau_1+t_2)x+\tau_2 y\} - \sin\{(\tau_1-t_2)x-\tau_2 y\}]$.

(3) $\sin(tx+\tau y)$
$+ 2(-1)^{m_1}(-1/\tau_1)[\sin\{(2t_1+t_2)x+\tau_2 y\} - \sin\{(2t_1-t_2)x-\tau_2 y\}]$.

(4) $(-1/\tau)\cos(tx+\tau y)\csc x$
$+2(-1/\tau_2)[\sin\{(\tau_1+t_2)x+\tau_2 y\}+\sin\{(\tau_1-t_2)x-\tau_2 y\}]$.

(5) $(-1/\tau)\cos(tx+\tau y)\sec x$
$+2(-1)^{n_1}(-1/\tau_1\tau_2)[\cos\{(\tau_1+t_2)x+\tau_2 y\}+\cos\{(\tau_1-t_2)x-\tau_2 y\}]$.

(6) $(-1/\tau)\cos(tx+\tau y)$
$+2(-1)^{n_1}(-1/\tau_1\tau_2)[\cos\{(2t_1+t_2)x+\tau_2 y\}+\cos\{(2t_1-t_2)x-\tau_2 y\}]$.

(7) $\sin tx\csc x-2[\cos\{(t_2+\tau_1)x+2t_1 y\}-\cos\{(t_2-\tau_1)x-2t_1 y\}]$.

(8) $(-1/m)\sin tx\csc x$
$-2(-1/m_2)[\cos\{(t_2+\tau_1)x+2t_1 y\}-\cos\{(t_2-\tau_1)x-2t_1 y\}]$.

(9) $\sin tx\csc x$
$-2(-1)^{n_1}[\cos\{(t_2+\tau_1)x+2t_1 y\}-\cos\{(t_2-\tau_1)x-2t_1 y\}]$.

(10) $(-1/m)\sin tx\csc x$
$-2(-1)^{n_1}(-1/m_2)[\cos\{(t_2+\tau_1)x+2t_1 y\}-\cos\{(t_2-\tau_1)x-2t_1 y\}]$.

(11) $(-1/\tau)\cos tx\sec x$
$+2(-1/\tau_1\tau_2)[\cos\{(t_2+\tau_1)x+2t_1 y\}+\cos\{(t_2-\tau_1)x-2t_1 y\}]$.

(12) $(-1/t)\cos tx\sec x$
$+2(-1)^{n_1}(-1/\tau_1 t_2)[\cos\{(t_2+\tau_1)x+2t_1 y\}+\cos\{(t_2-\tau_1)x-2t_1 y\}]$.

(13) $(-1/\tau)\cos(\tau x+ty)$
$+2(-1/\tau_1\tau_2)[\cos\{(2t_1+\tau_2)x+t_2 y\}+\cos\{(2t_1-\tau_2)x-t_2 y\}]$.

4. There are twelve expansions in which the partition is (II). All are of the form

$$\phi_{abc}(x,y)=T_{abc}(x,y)+4\sum_{(n)}q^n(\sum F_{abc}(x,y;n)),$$

in which $T_{abc}(x,y)$ is a trigonometric term independent of the partition, and the rest of the notation is obvious from § 3. As in § 3 it is sufficient to tabulate here only 8.

	abc	K_{abc}	$T_{abc}(x,y)$ (Partition (II))
(1)	011	$\vartheta'_1\vartheta_2\vartheta_3$	$\csc x\csc y$
(2)	021	$\vartheta'_1\vartheta_0\vartheta_2$	$\sec x\csc y$
(3)	031	$\vartheta'_1\vartheta_0\vartheta_3$	$\csc y$
(4)	312	$\vartheta'_1\vartheta_2\vartheta_3$	$\csc x\sec y$
(5)	322	$\vartheta'_1\vartheta_0\vartheta_2$	$\sec x\sec y$
(6)	332	$\vartheta'_1\vartheta_0\vartheta_3$	$\sec y$
(7)	311	$\vartheta'_1\vartheta_0\vartheta_2$	$\csc x\csc y$
(8)	022	$\vartheta'_1\vartheta_2\vartheta_3$	$\sec x\sec y$.

The corresponding F's follow.

(1) $\sin(2tx+\tau y)\csc x+\sin\tau x\csc y$
$-2[\cos\{(\tau_1+2t_2)x+\tau_2 y\}-\cos\{(\tau_1-2t_2)x-\tau_2 y\}]$.

(2) $(-1/\tau)(-1)^n\cos\tau x\csc y+\sin(2tx+\tau y)\sec x$
$+2(-1)^{n_1}(-1/\tau_1)[\sin\{(\tau_1+2t_2)x+\tau_2 y\}-\sin\{(\tau_1-2t_2)x-\tau_2 y\}]$.

(3) $\sin(2tx+\tau y)+(-1)^n(-1/\tau)\cos 2tx\csc y$
$+2(-1)^{n_1}(-1/\tau_1)[\sin\{2t_1+2t_2)x+\tau_2 y\}-\sin\{(2t_1-2t_2)x-\tau_2 y\}]$.

(4) $(-1/\tau)\cos(2tx+\tau y)\csc x+\sin\tau x\sec y$
$+2(-1/\tau_2)[\sin\{\tau_1+2t_2)x+\tau_2 y\}+\sin\{(\tau_1-2t_2)x-\tau_2 y\}]$.

(5) $(-1/\tau)[\cos(2tx+\tau y)\sec x+(-1)^n\cos\tau x\sec y]$
$+2(-1)^{n_1}(-1/\tau_1\tau_2)[\cos\{\tau_1+2t_2)x+\tau_2 y\}+\cos\{(\tau_1-2t_2)x-\tau_2 y\}]$.

(6) $(-1/\tau)[\cos(2tx+\tau y)+(-1)^n\cos 2tx\sec y]$
$+2(-1)^{n_1}(-1/\tau_1\tau_2)[\cos\{(2t_1+2t_2)x+\tau_2 y\}+\cos\{(2t_1-2t_2)x-\tau_2 y\}]$.

(7) $(-1)^n[\sin(2tx+\tau y)\csc x+\sin\tau x\csc y]$
$-2(-1)^n[\cos\{(\tau_1+2t_2)x+\tau_2 y\}-\cos\{(\tau_1-2t_2)x-\tau_2 y\}]$.

(8) $(-1/\tau)[(-1)^n\cos(2tx+\tau y)\sec x+\cos\tau x\sec y]$
$+2(-1)^{n_2}(-1/\tau_1\tau_2)[\cos\{(\tau_1+2t_2)x+\tau_2 y\}+\cos\{(\tau_1-2t_2)x-\tau_2 y\}]$.

5. There are 4 expansions in which the partition is (III). All are of the form

$$\phi_{abc}(x,y)=T_{abc}(x,y)+4\sum_{(m)}q^m\Big(\sum F_{abc}(x,y;m)\Big)+4\sum_{(n)}q^{2n}\Big(\sum G_{abc}(x,y;2n)\Big)$$

and as before we need list only 2.

	abc	K_{abc}	$T_{abc}(x,y)$ (Partition III)
(1)	020	$\vartheta'_1\vartheta_0\vartheta_3$	$\sec x$
(2)	310	$\vartheta'_1\vartheta_0\vartheta_3$	$\csc x$.

The corresponding F, G follow.

(1) $F_{020}(x,y;m)=\sin(\tau x+2ty)\tan x$
$+2(-1)^{d_1+\delta_1}[\cos\{(2d_1+\tau_2)x+2t_2y\}-\cos\{(2d_1-\tau_2)x-2t_2y\}]$;
$-G_{020}(x,y;2n)=(-1)^{d+\delta}\sin 2dx\csc x-\sin(\tau' x+2t'y)\tan x+$
$-2(-1)^{d'_1+\delta'_1}[\cos\{(2d'_1+\tau'_2)x+2t'_2y\}-\cos\{(2d'_1-\tau'_2)x-2t'_2y\}]$.

(2) $F_{310}(x,y;m)=(-1/\tau)\cos(\tau x+2ty)\cot x+$
$+2(-1)^{\delta_1}(-1/\tau_2)[\sin\{(2d_1+\tau_2)x+2t_2y\}+\sin\{(2d_1-\tau_2)x-2t_2y\}]$;
$G_{310}(x,y;2n)=(-1/\tau')\cos(\tau' x+2t'y)\cot x+(-1)^{\delta}\sin 2dx\sec x+$
$+2(-1)^{\delta'_1}(-1/\tau'_2)[\sin\{(2d'_1+\tau'_2)x+2t'_2y\}+\sin\{(2d'_1-\tau'_2)x-2t'_2y\}]$.

6. For partition (IV) there are 8 expansions, of which only 4 need be tabulated. All are of the form

$$\phi_{abc}(x, y) = 4 \sum_{(m)} q^{m/2} (\sum F_{abc}(x, y; m)).$$

	abc	K_{abc} (Partition IV)
(1)	202	$\vartheta'_1 \vartheta_2 \vartheta_3$.
(2)	131	$\vartheta'_1 \vartheta_0 \vartheta_2$.
(3)	132	$\vartheta'_1 \vartheta_2 \vartheta_3$.
(4)	201	$\vartheta'_1 \vartheta_0 \vartheta_2$.

The corresponding F's follow.

(1) $\sin tx(\cot x - \tan y)$
$-2(-1)^{\delta_2}[\cos\{(t_1 + 2d_2)x + 2\delta_2 y\} - \cos\{(t_1 - 2d_2)x - 2\delta_2 y\}]$.

(2) $(-1/m)\sin tx(\cot x + \cot y)$
$-2(-1/m_1)[\cos\{(t_1 + 2d_2)x + 2\delta_2 y\} - \cos\{(t_1 - 2d_2)x - 2\delta_2 y\}]$.

(3) $(-1/t)\cos tx(\tan x + \tan y)$
$-2(-1/t_1)(-1)^{d_2+\delta_2}[\sin\{(t_1 + 2d_2)x + 2\delta_2 y\} - \sin\{(t_1 - 2d_2)x - 2\delta_2 y\}]$.

(4) $(-1/\tau)\cos tx(-\tan x + \cot y)$
$+2(-1/\tau_1)(-1)^{d_2}[\sin\{(t_1 + 2d_2)x + 2\delta_2 y\} - \sin\{(t_1 - 2d_2)x - 2\delta_2 y\}]$.

7. For partition (V) there are 4 expansions, of which only 3 need be tabulated. All are of the form

$$\phi_{abc}(x, y) = T_{abc}(x, y) + 4 \sum_{(n)} q^{2n} (\sum F_{abc}(x, y; 2n)),$$

where $T_{abc}(x, y)$ is a trigonometric term independent of the partition.

	abc	K_{abc}	$T_{abc}(x, y)$ (Partition V)
(1)	222	$\vartheta'_1 \vartheta_0 \vartheta_3$	$1 - \tan x \tan y$
(2)	121	$\vartheta'_1 \vartheta_0 \vartheta_3$	$1 + \tan x \cot y$
(3)	211	$\vartheta'_1 \vartheta_0 \vartheta_3$	$-1 + \cot x \cot y$.

The corresponding F's follow.

(1) $(-1)^{\delta}[\sin 2(dx + \delta y)\tan x + (-1)^{d} \sin 2dx(\tan y - \cot x)] +$
$+ 2(-1)^{d_1+\delta_1+\delta_2}[\cos 2\{(d_1 + d_2)x + \delta_2 y\} - \cos 2\{(d_1 - d_2)x - \delta_2 y\}]$.

(2) $\sin 2(dx + \delta y)\tan x - (-1)^{d+\delta} \sin 2dx(\cot x + \cot y) +$
$+ 2(-1)^{d_1+\delta_1}[\cos 2\{(d_1 + d_2)x + \delta_2 y\} - \cos 2\{(d_1 - d_2)x - \delta_2 y\}]$.

(3) $(-1)^{\delta} \sin 2dx(\cot y - \tan x) + (-1)^{d} \sin 2(dx + \delta y)\cot x +$
$-2(-1)^{d_2+\delta_1}[\cos 2\{(d_1 + d_2)x + \delta_2 y\} - \cos 2\{(d_1 - d_2)x - \delta_2 y\}]$.

8. It will be sufficient to indicate how one of the foregoing expansions can be calculated. We have, for example,

$$\phi_{223}(x,y) = \vartheta'_1 \frac{\vartheta_2(x+y)}{\vartheta_0(x)\vartheta_3(y)} \cdot \vartheta_0\vartheta_2 \frac{\vartheta_0(x)}{\vartheta_2(x)} = \vartheta'_1\vartheta_0\vartheta_2\,\psi_{223}(x,y).$$

The expansions of the two functions which are multiplied are given in reduced form in the papers cited in § 1.† Note, however, that we also have,

$$\phi_{223}(x,y) = \vartheta'_1 \frac{\vartheta_2(x+y)}{\vartheta_2(x)\vartheta_1(y)} \cdot \vartheta_0\vartheta_2 \frac{\vartheta_1(y)}{\vartheta_3(y)} = \vartheta'_1\vartheta_0\vartheta_2\,\psi_{223}(x,y)\,;$$

this gives an expansion corresponding to partition (IV) whereas the preceding identity gives one belonging to partition (I). We are thus led to *arithmetically* different expansions for the same function, which is, of course, desirable from the point of view of the applications. Similarly all 48 expansions can be calculated. In what follows we list the expansions obtained by the second method of calculation; it will be sufficient to give the value of $F_{abc}(x,y;n)$ since K_{abc} and $T_{abc}(x,y)$ are the same as before.

9. The expansions corresponding to partition (I) with

$$\phi_{abc}(x,y) = 4\sum_{(m)} q^{m/2}\Big(\sum F_{abc}(x,y;m)\Big)$$

are as follows:

abc	$F_{abc}(x,y;m)$
130	$(-1/m)\sin(tx+\tau y)$ $+2(-1/\tau_1)(-1/m_2)[\sin\{t_2x+(\tau_2+2t_1)y\}+\sin\{t_2x+(\tau_2-2t_1)y\}]$.
233	$(-1/\tau)\cos(\tau x+ty)$ $+2(-1)^{m_1}(-1/\tau_1\tau_2)[\cos\{\tau_2x+(t_2+2t_1)y\}+\cos\{\tau_2x+(t_2-2t_1)y\}]$.
000	$\sin ty\csc y-2[\cos\{2t_1x+(\tau_1+t_2)y\}-\cos\{2t_1x+(\tau_1-t_2)y\}]$.
030	$(-1/\tau)\cos ty\sec y$ $+2(-1)^{m_1}(-1/\tau_1\tau_2)[\cos\{2t_1x+(\tau_1+t_2)y\}+\cos\{2t_1x+(\tau_1-t_2)y\}]$.
303	$(-1/t)\cos ty\sec y$ $+2(-1/\tau_1t_2)[\cos\{2t_1x+(\tau_1+t_2)y\}+\cos\{2t_1x+(\tau_1-t_2)y\}]$.
333	$(-1/m)\sin ty\csc y$ $-2(-1)^{m_1}(-1/m_2)[\cos\{2t_1x+(\tau_1+t_2)y\}-\cos\{2t_1x+(\tau_1-t_2)y\}]$.
202	$(-1/\tau)\cos(tx+\tau y)\sec y$ $+2(-1/\tau_1\tau_2)[\cos\{t_2x+(\tau_2+\tau_1)y\}+\cos\{t_2x+(\tau_2-\tau_1)y\}]$.

† See also, E. T. Bell, *Giornale di Matematiche*, vol. 61 (1921), pp. 93-114.

131 $(-1/m)\sin(tx+\tau y)\csc y$
$-2(-1)^{n_1}(-1/m_2)[\cos\{t_2x+(\tau_2+\tau_1)y\}-\cos\{t_2x+(\tau_2-\tau_1)y\}].$

300 $(-1/\tau)\cos ty\sec y$
$+2(-1/\tau_1\tau_2)[\cos\{2t_1x+(\tau_1+t_2)y\}+\cos\{2t_1x+(\tau_1-t_2)y\}].$

033 $(-1/t)\cos ty\sec y$
$+2(-1)^{n_1}(-1/\tau_1t_2)[\cos\{2t_1x+(\tau_1+t_2)y\}+\cos\{2t_1x+(\tau_1-t_2)y\}].$

132 $(-1/m)\sin(tx+\tau y)\sec y$
$+2(-1/\tau_1m_2)[\sin\{t_2x+(\tau_2+\tau_1)y\}+\sin\{t_2x+(\tau_2-\tau_1)y\}].$

201 $(-1/\tau)\cos(tx+\tau y)\csc y$
$+2(-1)^{n_1}(-1/\tau_2)[\sin\{t_2x+(\tau_2+\tau_1)y\}-\sin\{t_2x+(\tau_2-\tau_1)y\}].$

203 $(-1/\tau)\cos(tx+\tau y)$
$+2(-1/\tau_1\tau_2)[\cos\{t_2x+(\tau_2+2t_1)y\}+\cos\{t_2x+(\tau_2-2t_1)y\}].$

10. The expansions in this section correspond to partition (II) with

$$\phi_{abc}(x,y)=T_{abc}(x,y)+4\sum_{(n)}q^n(\sum F_{abc}(x,y;n)).$$

abc	$F_{abc}(x,y;n)$
011	$\sin(\tau x+2ty)\csc y+\sin\tau y\csc x$ $-2[\cos\{\tau_2x+(2t_2+\tau_1)y\}-\cos\{\tau_2x+(2t_2-\tau_1)y\}].$
021	$(-1)^n[(-1/\tau)\cos(\tau x+2ty)\csc y+\sin\tau y\sec x]$ $+2(-1)^n(-1/\tau_2)[\sin\{\tau_2x+(2t_2+\tau_1)y\}-\sin\{\tau_2x+(2t_2-\tau_1)y\}].$
312	$(-1)^n\sin(\tau x+2ty)\sec y+(-1/\tau)\cos\tau y\csc x$ $+2(-1)^{n_2}(-1/\tau_1)[\sin\{\tau_2x+(2t_2+\tau_1)y\}+\sin\{\tau_2x+(2t_2-\tau_1)y\}].$
322	$(-1/\tau)[\cos(\tau x+2ty)\sec y+(-1)^n\cos\tau y\sec x]$ $+2(-1)^{n_1}(-1/\tau_1\tau_2)[\cos\{\tau_2x+(2t_2+\tau_1)y\}+\cos\{\tau_2x+(2t_2-\tau_1)y\}].$
020	$(-1/\tau)[(-1)^n\cos(\tau x+2ty)+\cos 2ty\sec x]$ $+2(-1)^{n_2}(-1/\tau_1\tau_2)[\cos\{\tau_2x+(2t_2+\tau_1)y\}+\cos\{\tau_2x+(2t_2-\tau_1)y\}].$
310	$(-1)^n\sin(\tau x+2ty)+(-1/\tau)\cos ty\csc x$ $+2(-1)^{n_1}(-1/\tau_2)[\sin\{\tau_1x+(2t_1+2t_2)y\}+\sin\{\tau_1x+(2t_1-2t_2)y\}].$
311	$(-1)^n[\sin(\tau x+2ty)\csc y+\sin\tau y\csc x]$ $-2(-1)^n[\cos\{\tau_2x+(2t_2+\tau_1)y\}-\cos\{\tau_2x+(2t_2-\tau_1)y\}].$
022	$(-1/\tau)[(-1)^n\cos(\tau x+2ty)\sec y+\cos\tau y\sec x]$ $+2(-1)^{n_2}(-1/\tau_1\tau_2)[\cos\{\tau_2x+(2t_2+\tau_1)y\}+\cos\{\tau_2x+(2t_2-\tau_1)y\}].$

11. The expansions given below correspond to partition (III) with

$$\phi_{abc}(x,y)=T_{abc}(x,y)+4\sum_{(m)}q^m(\sum F_{abc}(x,y;m))$$
$$+4\sum_{(n)}q^{2n}(\sum G_{abc}(x,y;2n)).$$

We have for $(abc) = (031)$ or (332):

$$-F_{031}(x,y;m) = (-1/\tau)\cos(2tx+\tau y)\cot y$$
$$+2(-1)^{\delta_1}(-1/\tau)[\sin\{2t_2x+(\tau_2+2d_1)y\}-\sin\{2t_2x+(\tau_2-2d_1)y\}].$$
$$G_{031}(x,y;2n) = (-1)^{\delta}\sin 2dy\sec y+(-1/\tau')\cos(2t'x+\tau' y)\cot y+$$
$$+2(-1)^{\delta'_1}(-1/\tau'_2)[\sin\{2t'_2x+(2d'_1+\tau'_2)y-\sin\{2t'_2x+(\tau'_2-2d'_1)y\}].$$
$$-F_{332}(x,y;m) = \sin(2tx+\tau y)\tan y$$
$$+(-1)^{d_1+\delta_1}[\cos\{2t_2x+(\tau_2+2d_1)y\}-\cos\{2t_2x+(\tau_2-2d_1)y\}].$$
$$G_{332}(x,y;2n) = -(-1)^{d+\delta}\sin 2dy\csc y+\sin(2t'x+\tau' y)\tan y+$$
$$+2(-1)^{d'_1+\delta'_1}[\cos\{2t'_2x+(\tau'_2+2d'_1)y\}-\cos\{2t'_2x+(\tau'_2-2d'_1)y\}].$$

12. The following correspond to partition (IV) with

$$\phi_{abc}(x,y) = 4\sum_{(m)} q^{m/2}(\sum F_{abc}(x,y;m)).$$

abc	$F_{abc}(x,y;m)$
110	$(\cot x+\cot y)\sin ty$ $-2[\cos\{2d_2x+(2\delta_2+t_1)y\}-\cos\{2d_2x+(2\delta_2-t_1)y\}].$
120	$(-1/\tau)(\tan x+\tan y)\cos ty$ $-2(-1/\tau_1)(-1)^{d_2+\delta_2}[\sin\{2d_2x+(2\delta_2+t_1)y\}+\sin\{2d_2x+(2\delta_2-t_1)y\}].$
213	$(-1/t)(\cot x-\tan y)\cos ty$ $+2(-1/t_1)(-1)^{\delta_2}[\sin\{2d_2x+(2\delta_2+t_1)y\}+\sin\{2d_2x+(2\delta_2-t_1)y\}].$
223	$(-1/m)(-\tan x+\cot y)\sin ty$ $-2(-1/m_1)(-1)^{d_2}[\cos\{2d_2x+(2\delta_2+t_1)y\}-\cos\{2d_2x+(2\delta_2-t_1)y\}].$

13. Finally, the expansions corresponding to partition (V) with

$$\phi_{abc}(x,y) = T_{abc}(x,y)+4\sum_{(n)} q^{2n}(\sum F_{abc}(x,y;2n)),$$

are:

(abc)	$F_{abc}(x,y;2n)$
(222)	$(-1)^d[\sin 2(dx+\delta y)\tan y+(-1)^{\delta}\sin 2dy(\tan x-\cot y)]+$ $+2(-1)^{d_1+d_2+\delta_2}[\cos 2\{d_1x+(\delta_1+d_2)y\}-\cos 2\{d_1x+(\delta_1-d_2)y\}].$
(121)	$(-1)^{\delta}[-(-1)^{d}\sin 2(dx+\delta y)\cot y+\sin 2dy(\tan x+\tan y)]+$ $+2(-1)^{d_1+\delta_1+\delta_2}[\cos 2\{d_1x+(\delta_1+d_2)y\}-\cos 2\{d_1x+(\delta_1-d_2)y\}].$
(211)	$(-1)^{\delta}[\sin 2(dx+\delta y)\cot y+\sin 2dy(\cot x-\tan y)]+$ $-2(-1)^{\delta_1+\delta_2}[\cos 2\{d_1x+(\delta_1+d_2)y\}-\cos 2\{d_1x+(\delta_1-d_2)y\}].$

THE UNIVERSITY OF NEBRASKA.
CALIFORNIA INSTITUTE OF TECHNOLOGY.

CONCERNING CONTINUA OF FINITE DEGREE AND LOCAL SEPARATING POINTS.

By G. T. WHYBURN.

A point p of a compact metric continuum M is said to be a point of finite degree * of M provided that for each $\epsilon > 0$ there exists an uncountable family G of neighborhoods of p in M each of diameter $< \epsilon$ and each having a finite set for a boundary and such that for any two neiborhoods U, V of G we have either $\bar{U} \subset V$ or $\bar{V} \subset U$. Similarly, p is said to be of degree α provided α is the least cardinal number such that for each ϵ the neighborhoods of G can be chosen so that the boundary of each of them is of power $\leq \alpha$. (We shall be concerned principally with the case where α is an integer.) If X is a given subset of M, then p will be said to be of finite degree or of degree α in M *relative to* X provided that, in addition to satisfying the conditions previously stated, the neighborhoods of G can be so chosen that the boundary of each of them is a subset of X. These notions concerning the degree of a point should not be confused with the analogous but distinctly different notions concerning the *order* or *index* of a point as introduced by Menger and Urysohn. Both concepts will be used in this paper, but careful reading should enable one to avoid confusing them.

A continuum is said to be of finite degree provided each of its points is a point of finite degree. In this paper we shall show, among other results, that such continua may be identified with the continua previously studied by the author † which have the property that any subcontinuum contains uncountably many local separating points ‡ of the given continuum.

(1) THEOREM. *All save possibly a countable number of the local separating points of any continuum M are points of degree 2 of M.*§

* See H. Kamiya, *Tôhoku Mathematical Journal*, vol. 36 (1933), pp. 58-72.

† See my paper "Decompositions of continua by means of local separating points," *American Journal of Mathematics*, vol. 55 (1933), pp. 437-457.

‡ The point p of a continuum M is called a local separating point of M provided there exists a neighborhood R of p in M such that p separates $\bar{R}$ between some pair of points of the component of $\bar{R}$ containing p. See my paper "Local separating points of continua," *Monatshefte für Mathematik und Physik*, vol. 36 (1929), pp. 305-314.

§ This theorem is closely related to Theorem 9 of my paper "Local separating points of continua," *loc. cit.*; compare also Kamiya, *loc. cit.*, Theorem 18.

Suppose on the contrary, that there exists an uncountable set H of local separating points of M no one of which is a point of degree 2. Then applying Theorem 10 of my paper just referred to, we obtain a connected open subset R of M bounded by two points α and β of H and such that α and β are separated in $\bar{R}$ by each point of an uncountable subset E of H. Let E_0 be the set of all points of E which are condensation points of E. Then since each point of E is a cut point of $\bar{R}$ it follows by a theorem of the author's * that E_0 contains a subset S such that $E_0 - S$ is countable and each point of S is a point of order 2 in $\bar{R}$ relative to S.

Now let p be any point of S and let ϵ be any positive number. There exists an ϵ-neighborhood U of p in M bounded by two points x and y of S. Let K be the set of all points in U which separate x and y in $\bar{U}$. Then $K \supset S \cdot U$ and the set $K + x + y$ can be † ordered in a natural linear order from x to y. Then since x and y are points of order 2 relative to S and every point of S is a condensation point of S and since furthermore K is the sum of a G_δ-set and a countable set,‡ it follows that there exist 0-dimensional perfect sets A and B in K such that if $a \in A$ and $b \in B$, we have the order $x < a < p < b < y$. Clearly there exists a $(1-1)$ order reversing correspondence (a, b) between A and B. Then each pair of corresponding points (a, b) determines an ϵ-neighborhood $U(a, b)$ in M bounded by $a + b$; and if (a_1, b_1) and (a_2, b_2) are any two such pairs and $a_1 < a_2$, we have $U(a_1, b_1) \supset \overline{U(a_2, b_2)}$. Thus p is of degree 2 in M, contrary to supposition; and our theorem is proven.

(1.1) *Any point of finite degree or of degree n is of the same degree in M relative to the local separating points of M of degree 2.*

For if p is of finite degree (or of degree n) in M, then for each $\epsilon > 0$ we have an uncountable family G of ϵ-neighborhoods satisfying the conditions of the definition and such that the boundary of each neighborhood is finite (or of power $\leq n$). Now if we choose the neighborhoods U of G so that every point of the boundary of U is a limit point of $M - U$, as we clearly can do, then every boundary point of a neighborhood of G is a local separating point of M. Then since by (1) all save a countable number of the local separating points of M are points of degree 2, it follows that for all save a

* See my paper in *Transactions of the American Mathematical Society*, vol. 33 (1931), p. 450.

† See my paper, *loc. cit.*, p. 446.

‡ See my paper in *Transactions of the American Mathematical Society*, vol. 32 (1930), p. 151.

countable number of the neighborhoods U of G, $F(U)$ consists wholly of local separating points of M of degree 2; and our result follows at once.

Throughout the rest of the paper, if M is a continuum then we shall denote by L the set of all local separating points of M and by Q the set of all points of L which are of degree 2 in M. Now from (1) and (1.1) we have at once

(1.2) *Every point Q is of degree 2 in M relative to Q and $L - Q$ is countable.*

Now since every countable set which separates a continuum M contains at least one local separating point of M,* a slight modification of the argument given for (1.1) suffices to prove

(1.3) *Every point of countable degree* (i. e., $\leq \aleph_0$) *of a continuum M is a condensation point of local separating points of degree 2 of M.*

(2) THEOREM. *If p is a point of finite order or of order n in a compact continuum M relative to Q, then p is of the same degree in M relative to Q.*

Proof. Given $\epsilon > 0$, there exists an $\epsilon/3$-neighborhood R of p whose boundary consists of a finite number, k, of points $p_1, p_2, \cdots, p_k$ of Q where $k = n$ in case p is of order n relative to Q. For each $i \leq k$ there exists an $\epsilon/3$-neighborhood V_i of p_i whose boundary consists of just two points x_i and y_i one of which, x_i, lies in R and the other in $M - \bar{R}$, and these can be chosen so that $\bar{V}_i \cdot \bar{V}_j = 0$, $i \neq j$, and $\bar{V}_i \cdot p = 0$. Now since for each $i \leq k$, p_i is of degree 2, it follows that if K_i denotes the set of all points separating y_i and p_i in $\bar{V}_i$, then K_i is uncountable and, in fact, contains a 0-dimensional perfect set A_i [see the proof of (1)]. The sets K_i are ordered from p_i to y_i as in (1); and clearly there exists a $(1-1-1-\cdots-1)$ order preserving correspondence $(a_1, a_2, a_3, \cdots, a_k)$ between the sets $A_1, A_2, \cdots, A_k$, i. e., a grouping of the points of these sets into groups $(a_1, a_2, \cdots, a_k)$, $a_i \in A_i$ such that when $a_1^1, a_1^2 \in A_1$ and $a_1^1 < a_1^2$ in K_1, then for each $i \leq k$, $a_i^1 < a_i^2$ in K_i. It is seen at once that each such group $(a_1, a_2, \cdots, a_k)$ is the boundary of an ϵ-neighborhood $U(a_1, \cdots, a_k)$ of p obtained by adding to R certain points of the neighborhoods V_i and that $a_1^1 < a_1^2$, $(a_1^1, a_1^2 \in A_1)$, implies $\overline{U(a_1^1, a_2^1, \cdots, a_k^1)} \subset U(a_1^2, a_2^2, \cdots, a_k^2)$. Thus p is of the same degree in M as its order in M relative to Q; and hence, by (1.1), it is likewise of the same degree in M relative to Q.

* See my paper "Local separating points of continua," *loc. cit.*

(2.1) COROLLARY. *The points of finite degree or of degree n of a continuum M are exactly the points of finite order or of order n in M relative to Q.*

(2.2) (Kamiya, *loc. cit.*) *The points of M of finite degree or of degree $\leq n$ in M form a G_δ-set.*

(2.3) *Any point of a compact continuum M of infinite degree lies in a non-degenerate continuum of such points.**

This is an immediate consequence of (2.1) and of a theorem of Menger's.†

(2.4) *In order that a compact continuum M be of finite degree it is necessary and sufficient that any two points of M be separated in M by a finite number of points of Q.*

This can be proven at once from (2) by a simple application of the Borel Theorem.

(3) THEOREM. *In order that a point p be a point of finite degree of a compact continuum M it is necessary and sufficient that $\dim_p(M - Q + p) = 0$.*

The necessity of the condition follows at once from (1.2). To prove the sufficiency, let us suppose $\dim_p(M - Q + p) = 0$ and $\epsilon > 0$. Then there exists an $\epsilon/3$-neighborhood R of p with $F(R) \subset Q$. Now for each $x \in F(R)$, it follows by (1.3) that there exists an $\epsilon/3$-neighborhood U_x of x such that $F(U_x)$ consists of exactly two points of Q. Applying the Borel Theorem we obtain a finite number of the neighborhoods U_x, say $U_1, U_2, \cdots, U_k$ whose sum covers $F(R)$. Then if $U = R + U_1 + U_2 + \cdots + U_n$, it is clear that U is an ϵ-neighborhood of p whose boundary consists of n or less points of Q. Thus p is of finite order in M relative to Q, and hence by (2) it is also of finite degree in M.

(3.1) COROLLARY. *In order that a continuum M be of finite degree it is necessary and sufficient that $\dim(M - Q) = 0$.*

Note. It follows at once from this that if M^2 denotes the set of all points of M of order 2, then $\dim(M - M^2) = 0$, because $M - M^2 \subset M - Q$. (See Kamiya, *loc. cit.*) However, examples are easily constructed to show that the condition that $\dim(M - M^2) = 0$ is not sufficient to make M of finite degree.

(4) THEOREM. *In order that a compact continuum M be of finite degree*

* Compare with Kamiya, *loc. cit.*, Theorem 10.

† See *Kurventheorie*, Teubner, 1928, p. 128.

at a point p it is necessary and sufficient that every subcontinuum of M containing p contain uncountably many local separating points of M.

The condition is sufficient. For suppose the condition is satisfied at p. Then if p were not a point of finite degree, by (2.2) there would exist a non-degenerate subcontinuum N of M containing p no point of which is of finite degree. But this is impossible, since by hypothesis N must contain uncountably many local separating points of M; and hence, by (1), it contains at least one point of degree 2.

The condition is also necessary. For let p be of finite degree and let N be any subcontinuum of M containing p. Then there exists an uncountable family $[U]$ of neighborhoods of p of the type we have been considering such that for any $U \in [U]$, $\delta(U) < \delta(N)$ and $F(U)$ is a finite set of local separating points of M. But clearly for each U we have $F(U) \cdot N \neq 0$, and hence N contains uncountably many local separating points of M.

(4.1) COROLLARY. *In order that a compact continuum M be of finite degree it is necessary and sufficient that every subcontinuum of M contain uncountably many local separating points of M.*

Now the property involved in this corollary is readily seen to be cyclicly extensible and reducible.* Hence we have

(4.2) *The property of being a continuum of finite degree is cyclicly extensible and reducible.*

Following the notation used in my paper "Decompositions of continua by means of local separating opints" (*loc. cit.*), for any point p of a continuum M let us denote by $C_1(p)$ the maximal subcontinuum of M containing p and containing only a countable number of local separating points of M. Then by (4) we have

(4.3) *In order that p be a point of finite degree of a compact continuum M it is necessary and sufficient that $C_1(p) = p$.*

Now it was shown in the author's paper just mentioned that the decomposition of M into the sets $[C_1(p)]$ is upper semi-continuous † and that the hyperspace C_1 of this decomposition has the property that any sub-

* See Kuratowski and Whyburn, *Fundamenta Mathematicae*, vol. 16 (1930), pp. 305-331.

† See R. L. Moore, "Foundations of point set theory," *American Mathematical Society Colloquium Publications*, 1932, Chapter V.

continuum of C_1 contains uncountably many local separating points of C_1. Thus by (4.1), C_1 is of finite degree. Thus we have

(4.4) *Every compact continuum M admits of an upper semi-continuous decomposition into continua with a continuum of finite degree as hyperspace.*

(The hyperspace reduces to a single point if and only if the local separating points of M are countable).

(5) THEOREM. *If M is a regular curve of order ≤ 3, then degree and order are identical for all points of M.*

For let p be a point of such a curve M, let k be the order of p and let $\epsilon > 0$. Then there exists an $\epsilon/3$-neighborhood R of p whose boundary consists of k points $p_1, \cdots, p_k$, $(k \leq 3)$, each point of which is a limit point of $M - R$. For each $i \leq k$, p_i is a point of order 1 either of $\bar{R}$ or of $\overline{M - R}$, say of $\bar{S}$. Then there exists an $\epsilon/3$-neighborhood V_i of p_i whose boundary contains just one point q_i of $\bar{S}$ and such that $\bar{V}_i \cdot \bar{V}_j = 0$, $i \neq j$ and $\bar{V}_i \cdot p = 0$. Now since no two true cyclic elements of $\bar{V}_i$ can have a common point * and since p_i is a point of order 1 of $\bar{S}$, it follows at once that p_i and q_i are separated in $\bar{V}_i$ by uncountably many distinct points. Since all such points are local separating points of M, it results from (1) that p_i and q_i are separated in $\bar{V}_i$ by at least one point x_i of degree 2 (i. e., a point of the set Q). It is easily seen then that the points $x_1, \cdots, x_k$ bound an ϵ-neighborhood U of p. Thus p is of order k in M relative to Q and accordingly, by (2), p is of degree k.

In contrast to the theorem just proven, it is notable that *there exist regular curves of order 4 which are of degree $\mathfrak{c}$ at every point,* where $\mathfrak{c}$ denotes the power of the continuum. For the Sierpinski triangle curve † is an example of such a curve. To see that every point of this curve is of degree $\mathfrak{c}$, in view of (1.3) we have only to note that its local separating points are countable.

THE UNIVERSITY OF VIRGINIA.

* See my paper in the *Bulletin of the American Mathematical Society*, vol. 35 (1929), p. 223.

† See *Comptes Rendus*, vol. 162, p. 629.

THE TOPOLOGY OF (PATH) SURFACES.

By CHARLES B. MORREY, JR.†

In the author's recent work on the analytic characterization of surfaces of finite area, a considerable number of difficulties were encountered, due to the complicated topological structure of the most general surface. The object of the present paper is to study this structure and to provide a systematic treatment of a number of concepts and results which facilitate the discussion of the general surface.

Before giving a definition of what we mean by a (path) surface, we must introduce the notation. We shall use the ordinary notations for the sum and product of point sets. If S is any set, we shall designate by $\bar{S}$ the set S plus its limit points and by S^* the frontier points of S. If P is a point and S a set, we write $P \in S$ instead of "P belongs to S" and $P \notin S$ instead of "P does not belong to S." We write $\Sigma \subset S$ instead of Σ is a subset of S. By a Jordan region r (in n-space) we mean the interior of a simple closed $(n-1)$-manifold; we shall frequently also refer to $\bar{r}$ as a Jordan region. We designate the interior of a circle with center at P and radius r by $C(P, r)$.

Further, we shall use a vector notation throughout: the letters u and U will stand for the coördinate vector of a point in a set on which a surface is represented and the letters x and X for that of a point in the (N dimensional) space in which the surface lies. $U(u)$, $x(u)$, $X(U)$, etc. will be vector functions in these spaces; $U_1 \pm U_2$, $x_1 \pm x_2$, etc., will denote the sum or difference vectors, and $|\phi|$ the length of the vector ϕ. If P is in the U space, for instance, U_P will denote the U-coördinate vector of P. The distance between two points P and Q will be denoted by $d(P, Q)$ or by $|u_P - u_Q|$ or $|U_P - U_Q|$ if P and Q are in the u or U space respectively. *All vector functions will be assumed to be continuous.*

Now, let $x_1(u)$ and $X_2(U)$, $u \in \bar{r}$, $U \in \bar{R}$ (Jordan regions in n-space), be two vector functions. Let T, $T : U = U(u)$, be a $1-1$ continuous (sense preserving if desired) transformation of $\bar{r}$ into $\bar{R}$. Define $x_2(u) = X_2[U(u)]$ and $D_T(x_1, X_2)$ as the maximum of $|x_1(u) - x_2(u)|$ for $u \in \bar{r}$. Then we define $\| x_1, X_2 \|$ as the greatest lower bound of $D_T(x_1, X_2)$ for all T. By a (*path*) *manifold, M,* we shall mean a class of equations $x = x(u)$, such that if $x_1(u)$, and $X_2(U)$ are any two of the $x(u)$, then $\| x_1, X_2 \| = 0$. Any one of the equations of the class will be called a representation of the manifold.

† National Research Fellow (1931-1933).

It is easy to see that if $x_1(u)$, $X_2(U)$, and $\hat{X}_3(\hat{U})$, $u \in \bar{r}$, $U \in \bar{R}$, $\hat{U} \in \bar{\hat{R}}$, then $\| x_1, \hat{X}_3 \| \leq \| x_1, X_2 \| + \| X_2, \hat{X}_3 \|$, and also that $\| x_1, X_2 \| = \| X_2, x_1 \|$. Thus if $\| x_1, X_2 \| = \| X_2, \hat{X}_3 \| = 0$, then $\| x_1, \hat{X}_3 \| = 0$. If $n = 1$, the manifold will be called a *curve,* and if $n = 2$, it will be called a *surface.* It is clear that a manifold is defined by any one of its representations. This definition of manifold is clearly the logical consequence of Fréchet's † definition of the *distance between two surfaces* (or manifolds). This may be defined as follows: Let S_1 be defined by the equation $x = x_1(u)$, $u \in \bar{r}$, and S_2 by $x = X_2(U)$, $U \in \bar{R}$; then the Fréchet distance $\| S_1, S_2 \| = \| x_1, X_2 \|$, it being clear from the above that this is independent of the particular representations. Clearly $\| S_1, S_2 \| = 0$ when and only when S_1 and S_2 are identical.

I shall now merely give information enabling one to define a point of a surface. Let S, $S : x = x(u)$, $u \in \bar{r}$, be a surface and let $u_1 \neq u_2$ in $\bar{r}$. Then the point P_1 of S defined by $x = x(u_1)$ will coincide with P_2, that defined by $x = x(u_2)$, if and only if $x(u)$ is constant over a continuum of $\bar{r}$ containing u_1 and u_2. The fact that (see § 2) the collection of maximal continua over each of which $x(u)$ is constant is an "upper semicontinuous collection (see § 2) of mutually exclusive continua filling up $\bar{r}$," suggests an intimate relation between the structure of a surface and the theory of these collections developed by R. L. Moore. With his definition of limit element in such a collection, it is seen that there is a $1 - 1$ continuous correspondence between the continua of the above collection and the points of S. Furthermore, it will appear later that if $x = x(u)$, $u \in \bar{r}$, and $x = X(U)$, $U \in \bar{R}$, are representations of the same surface, then there is a $1 - 1$ continuous correspondence between the collections of continua in $\bar{r}$ and $\bar{R}$, each one of which corresponds to a point of S, so that the definition of identity or distinctness of two points of S is independent of the particular representation.

By a non-degenerate surface, we shall mean one which possesses a representation $x = x(u)$, $u \in \bar{r}$, in which $x(u)$ is not constant over any continuum of $\bar{r}$ (i. e. the points of S are in a $1 - 1$ continuous correspondence with those of $\bar{r}$). Unfortunately not all surfaces are non-degenerate. To study the structure of the general surface, the notion of a "hemicactoid" is introduced. It is shown that if $\bar{H}$ is a hemicactoid and $X(U)$ is a (continuous) vector function defined on it, then (1) there exists a "monotone" transformation (carrying only continua into points) $U = U(u)$, $u \in \bar{r}$, of $\bar{r}$ into $\bar{H}$ (which transformation also satisfies another interesting condition) and (2) if we define $x(u) = X[U(u)]$, then all the surfaces $x = x(u)$ thus obtained are

† M. Fréchet, "Sur la distance de deux surfaces," *Annales de la Société Polonaise de Mathématiques,* vol. 3 (1924), pp. 4-19.

identical. We then introduce parametric representations of surfaces on these hemicactoids and it is shown that every surface can be represented on some hemicactoid $\bar{H}$ in such a way that the representing vector function is not constant over any continuum of $\bar{H}$. Some useful theorems about the relations between two different representations of the same surface are demonstrated. Finally, a criterion, in terms of its given representation, that a surface be non-degenerate is developed and it is shown that a surface "bounded" by a Jordan curve and represented by a vector function, all of whose components are monotone in the sense of Lebesgue, is non-degenerate.

1. *On continuous curves in general.* In this section we shall recall a number of general definitions and results about continuous curves and shall develop a representation of a continuous curve which will be of use in later work. This representation is entirely analogous to a representation of an acyclic (see Definition 14, below) continuous curve developed by R. L. Wilder,† the arcs of his representation being replaced, in the present one, by simple cyclic chains (see below) of the continuous curve. All point sets considered in this section will be supposed to be in a bounded portion of Euclidean n-space.

Definitions 1-16 and Lemmas 1-7 are all either to be found in the literature or are easily deducible from known results. They are merely included to furnish an adequate introduction to the terminology and certain of the results of modern point set theory.

Definition 1. By a *maximal* set possessing a given property, we shall mean a set which is not a proper subset of any other set possessing the given property.

Definition 2. By a *component* of a point set, S, we shall mean a maximal connected subset of S (i. e. a maximal set possessing the property of being a connected subset of S).

Definition 3. A set is said to be *completely disconnected* if all of its components are points.

Definition 4. A point P of a connected set S is said to be a *cut point* of S if $S - P$ is not connected. Two sets H and K are said to be *mutually separated* if they have no points in common and neither contains a limit point of the other. If H, K, and T are proper subsets of the connected point set M, then T is said to *separate H from K in M* if $M - T$ is the sum of two mutually

† R. L. Wilder, "Concerning continuous curves," *Fundamenta Mathematicae*, vol. 7 (1925), p. 365.

separated point sets containing H and K respectively. In particular, H, K, and T may be points.

Definition 5.† ‡ A point, P, of a continuum $\bar{C}$, is said to be an *end point* of $\bar{C}$ provided that if $\bar{c}$ is a subcontinuum of $\bar{C}$ containing P, then P is not a limit point of any connected subset of $\bar{C} - \bar{c}$.

Definition 6. A point set D, is said to be *open with respect to a set* S, if $D = S \cdot G$, where G is an open subset of the given n-space.

Definition 7.§ The point set S, is said to be *connected in kleinen* at the point P, if $P \in S$ and every subset D, of S, which is open in S and contains P contains a subset d, having all these properties and which is a subset of a component of D.

Definition 8.§ The point set S, is said to be *locally connected* at the point P, if $P \in S$ and every subset D, of S, which is open in S and contains P contains a connected subset d, with all of these properties.

Definition 9.§ A *continuous curve* is a continuum which is connected in kleinen at each of its points.

LEMMA 1. *Any two points of a connected subset D, of a continuous curve $\bar{M}$, which is open in $\bar{M}$ may be joined by a simple arc which is a subset of D.*

Proof. This is a subcase of Theorem 10, chapter II, P. S. T.

LEMMA 2. *A necessary and sufficient condition that a bounded continuum $\bar{M}$, be a continuous curve is that it should be uniformly arcwise connected in kleinen, i. e. for every $\epsilon > 0$, there exists a $\delta > 0$ such that if P and Q are any two points of $\bar{M}$ at a distance $< \delta$ apart, they can be joined by an arc of $\bar{M}$ of diameter $< \epsilon$.*

Proof. This theorem is well known. The necessity of the condition follows easily from the above lemma and the sufficiency from the Definitions 7, 8, and 9.

LEMMA 3. *A necessary and sufficient condition that a bounded continuum $\bar{M}$ be a continuous curve is that there exists a continuous vector*

---

† G. T. Whyburn, "Concerning the structure of a continuous curve," *American Journal of Mathematics*, vol. 50 (1928), pp. 167-194, § 1. This paper will hereafter be referred to as W.

‡ G. T. Whyburn, "Concerning continua in the plane," *Transactions of the American Mathematical Society*, vol. 29 (1929) p. 382.

§ R. L. Moore, "Foundations of point set theory," *Colloquium Publications*, vol. 13, p. 94. We shall hereafter refer to this book as P. S. T.

function $x(u)$, $u \in \bar{r}$, $\bar{r}$ *being a Jordan region in k-space, such that the transformation* $x = x(u)$, *carries* $\bar{r}$ *into* $\bar{M}$.

Proof. This theorem was originally proved for the case that $\bar{r}$ is an interval.† It is easy to see that any k dimensional Jordan region is a continuous curve, and also that there exists a continuous transformation of such a region into an interval. Hence the above theorem follows.

Definition 10 (P. S. T., p. 63 and p. 72). If $\bar{M}$ is a continuum and P is a point of $\bar{M}$ such that there do not exist two points A and B such that (1) P separates A and B in $\bar{M}$ and (2) P is the only point separating A and B in $\bar{M}$, then P is a proper point of $\bar{M}$ and the set of all points X of $\bar{M}$ which are not separated from P in $\bar{M}$ by a point of $\bar{M}$ is called a *simple link* of $\bar{M}$. A simple link containing two points will be called *non-degenerate.*

Definition 11 (W., § 1). A continuous curve $\bar{M}$, is said to be *cyclically connected* if every two of its points lie on a simple closed curve of $\bar{M}$. A *maximal cyclic curve of* $\bar{M}$ is a maximal set $\bar{C}$, possessing the properties: (1) $\bar{C}$ is a cyclically connected continuous curve and (2) $\bar{C}$ is a subset of $\bar{M}$.

Definition 12 (W., § 1). By an *internal point* of a maximal cyclic curve $\bar{C}$, of $\bar{M}$ is meant a point of $\bar{C}$ which is not a cut point of $\bar{M}$.

Definition 13 (W., § 1). A subset E, of $\bar{M}$ is called a *cyclic element* of $\bar{M}$ if it belongs to one of the following three classes: (a) maximal cyclic curves of $\bar{M}$, (b) cut points of $\bar{M}$, and (c) end points of $\bar{M}$. Those of class (a) constitute the *non-degenerate cyclic elements of* $\bar{M}$.

LEMMA 4 (P. S. T., chapter II, Theorems 68-70). *Every simple link of a continuous curve* $\bar{M}$, *is a cyclic element of* $\bar{M}$. *Every non-degenerate cyclic element of* $\bar{M}$, *is a simple link of* $\bar{M}$. *A necessary and sufficient condition that a degenerate cyclic element* P, *of* $\bar{M}$ *be a simple link of* $\bar{M}$ *is that it does not belong to any non-degenerate simple link of* $\bar{M}$.

LEMMA 5.‡ *Every continuous curve is the sum of its cyclic elements.*

LEMMA 6 (W., § 1). (1) *No two cyclic elements of* $\bar{M}$ *have more than one point in common,* (2) *the common part of every pair of cyclic elements of* $\bar{M}$ *is either vacuous or itself a cyclic element of class* (b), *and* (3) *the set of non-internal points of a maximal cyclic curve of* $\bar{M}$ *is countable.*

---

† H. Hahn, "Über die allgemeinste ebene Punktmenge, die stetiges Bild einer Strecke ist," *Jahresbericht der Deutschen Mathematiker Vereinigung*, vol. 23 (1914), pp. 318-322.

‡ G. T. Whyburn, "Cyclically connected continuous curves," *Proceedings of the National Academy of Sciences*, vol. 13 (1927), pp. 31-38.

Definition 14 (P. S. T., p. 115). A continuous curve is said to be *acyclic* if it contains no simple closed curve.

Definition 15 (W., § 2). A point set X will be called a *simple cyclic chain* of $\bar{M}$ between two cyclic elements E_1 and E_2 of $\bar{M}$ provided that X is connected, contains E_1 and E_2, is the sum of some collection of the cyclic elements of $\bar{M}$, and is such that no proper connected subset of X contains E_1 and E_2 and is the sum of the elements of such a collection. The elements E_1 and E_2 are said to be the *end elements* of the chain.

LEMMA 7.† *If A and B are any two points of a continuous curve $\bar{M}$, K denotes the set of all those points of $\bar{M}$ which separate A from B in $\bar{M}$, and t is any arc in $\bar{M}$ from A to B, then* (1) *$K + A + B$ is a closed set of points of t and* (2) *each maximal segment S of $t - (K + A + B)$ determines a unique maximal cyclic curve of $\bar{M}$ which contains S.*

Definition 16. X is a *maximal simple cyclic chain* of $\bar{M}$ if it is not a proper subset of any other simple cyclic chain of $\bar{M}$.

LEMMA 8.‡ *Let $\bar{M}$ be a continuous curve in k-space and $\bar{C}$ a simple cyclic chain of $\bar{M}$. Then the point set $\bar{M} - \bar{C}$ is the sum of a finite or denumerable set of components M_n, where* (1) *$\bar{M}_n$ is a continuous curve,* (2) *$\bar{M}_n \cdot \bar{C} = \bar{M}_n - M_n =$ a single point,* (3) *$\bar{M}_m \cdot \bar{M}_n$, $m \neq n$, is null or a single point of $\bar{C}$,* (4) *every cyclic element of $\bar{M}$ containing a point of M_n lies in $\bar{M}_n$,* (5) *every arc joining a point of $\bar{M}_n$ to a point of $\bar{M} - M_n$ contains the point $\bar{M}_n - M_n$, and* (6) *for each $\epsilon > 0$, there are at most a finite number of the M_n of diameter $\geq \epsilon$.*

Proof. First of all, let M_n be one of the components of $\bar{M} - \bar{C}$. If M_m contained any limit points of M_n, $M_m + M_n$ would be connected and thus $m = n$. Hence $\bar{M}_n = M_n + \bar{M}_n \cdot \bar{C}$.

Now, let M'_n be the sum of all the cyclic elements of $\bar{M}$ which contain a point of M_n. Each point Q of $M'_n - M_n$ lies in a non-degenerate cyclic element of $\bar{M}$ containing a point P of M_n. Obviously no such point Q can lie in M_m, $m \neq n$, since each cyclic element of $\bar{M}$ is connected; thus every point (if any) of $M'_n - M_n$ lies in $\bar{C}$. Suppose Q is such a point; let c_P be the cyclic element of M'_n containing Q, and c_Q one in $\bar{C}$ containing it. Then

† G. T. Whyburn, "Some properties of continuous curves," *Bulletin of the American Mathematical Society*, vol. 33 (1927), pp. 305-308.

‡ The results of this lemma are contained in Theorems (4.3), (4.4), (4.7), and (6.5) of Kuratowski and Whyburn, "Sur les éléments cycliques, etc.," *Fundamenta Mathematicae*, vol. 16 (1930), pp. 305-331.

$c_P \neq c_Q$ so, from Lemma 6 we conclude that $Q = c_P \cdot c_Q$, that Q is a cut point of $\bar{M}$ and is thus the only point of $\bar{C} \cdot c_P$. Hence every point of $M'_n - M_n$ is a limit point of M_n and $\bar{M}'_n = \bar{M}_n$. This demonstrates (4).

Clearly M'_n is a connected set composed of cyclic elements of $\bar{M}$, so that by W., Theorems 8 and 11, $\bar{M}_n$ is a continuous curve (which demonstrates (1)) and contains all the simple arcs of $\bar{M}$ joining two points of $\bar{M}_n$. It is thus clear that $\bar{M}_n \cdot \bar{C}$ contains at most one point since $\bar{C}$ also possesses these properties. Clearly $\bar{M}_n \cdot \bar{C} \neq 0$ since any point of M_n can be joined to a point of $\bar{C}$ by an arc of $\bar{M}$ which has a last point in $\bar{M}_n$. Thus $\bar{M}_n \cdot \bar{C} = \bar{M}_n - M_n$ = a single point. From this and the previous sentence, (2), (3), and (5) follow.

Now we have seen that any arc of $\bar{M}$ joining a point of M_n and a point of $\bar{M} - \bar{M}_n$ must contain the point $Q_n = \bar{M}_n - M_n$. In each M_n, let P_n be as far from Q_n as possible. Now suppose there is an infinite subsequence of $\{M_n\}$, each of which is of diameter $\geq \epsilon$. Let $\{M_{n_k}\}$ be a subsequence of this so that the corresponding subsequence $\{P_{n_k}\}$ converges to P_0. Clearly $d(P_{n_k}, Q_{n_k}) \geq \epsilon/2$ so that there exists a K such that for $k > K$, $P_{n_k} \varepsilon\, C(P_0, \epsilon/4)$ and $Q_{n_k} \nvarepsilon\, C(P_0, \epsilon/4)$. Thus none of the P_{n_k}, $k > K$, can be joined to P_0 by an arc of $\bar{M}$ lying entirely in $C(P_0, \epsilon/4)$ since each such arc must contain Q_{n_k}. But this contradicts Lemma 2.

THEOREM 1.† *Every continuous curve $\bar{M}$ (in n-space) can be represented as the sum of a set M, described below, and a completely disconnected set M' of limit points of M, all of which are end points of $\bar{M}$. The set M is represented in the form $M = \bar{C}_1 + \bar{C}_2 \cdots$ where* (1) *each $\bar{C}_n$ is a simple cyclic chain of $\bar{M}$,* (2) *for each n, $\bar{M}_n = \bar{C}_1 + \cdots + \bar{C}_n$ is a continuous curve and $\bar{C}_{n+1} \cdot \bar{M}_n$ is a single cut point of $\bar{M}$ which belongs to an end cyclic element of $\bar{C}_{n+1}$ but not to an end cyclic element of $\bar{C}_1$, and* (3) *for each $\epsilon > 0$ there exists an $N(\epsilon)$ such that, for $n > N(\epsilon)$, the diameter of each $\bar{C}_n$ and of each component of $\bar{M} - \bar{M}_{n-1}$ is less than ϵ.*

Proof. Let d be the diameter of $\bar{M}$, P and Q be two points of $\bar{M}$ such that $d(P, Q) = d$, E and F cyclic elements of $\bar{M}$ containing P and Q respectively, and $\bar{\bar{C}}$ a maximal simple cyclic chain containing E and F such a chain being known to exist by W Theorems 13 and 15. Let $\{K_{n_1}\}$ be the sequence of components of $\bar{M} - \bar{\bar{C}}$. In each K_{n_1}, let P_{n_1} be a point as far from $Q_{n_1} = K_{n_1} \cdot \bar{\bar{C}}$ as possible, let E_{n_1} be a cyclic element of $\bar{M}$ containing P_{n_1}

† The essence of this theorem is to be found in Kuratowski and Whyburn, *loc. cit.*, § 8.

(this element lies in $\bar{K}_{n_1}$ by Lemma 8), let $F_{n_1} = Q_{n_1}$, and let $\bar{\bar{C}}_{n_1}$ be a simple cyclic chain of $\bar{K}_{n_1}$ containing E_{n_1} and F_{n_1} as endelements. Now let $\{K_{n_1,n_2}\}$ be the components of $\bar{K}_{n_1} - \bar{\bar{C}}_{n_1}$ for each n_1; the totality of the K_{n_1,n_2} are the components of $\bar{M} - \bar{\bar{C}} - \Sigma\bar{\bar{C}}_{n_1}$, since $K_{n_1,n_2} \subset K_{n_1}$. Also $\bar{K}_{n_1,n_2} \cdot \bar{K}_{n_1,n'_2}$ is null or a single point of $\bar{\bar{C}}_{n_1}$, $n'_2 \neq n_2$, and $\bar{K}_{n_1,n_2} \cdot \bar{K}_{n'_1,n'_2} = 0$, $n'_1 \neq n_1$. Let $Q_{n_1,n_2} = \bar{\bar{C}}_{n_1} \cdot \bar{K}_{n_1,n_2}$, let P_{n_1,n_2} be a point of $\bar{K}_{n_1,n_2}$ as far as possible from Q_{n_1,n_2} and determine a simple cyclic chain $\bar{\bar{C}}_{n_1,n_2}$ in $\bar{K}_{n_1,n_2}$ as above. Continue this indefinitely (i. e. unless $\bar{M} = \bar{M}_n$ for some n). If we order all of these $\bar{\bar{C}}_{n_1,\ldots,n_k}$ into a single sequence $\{\bar{C}_n\}$ so that $\bar{C}_n \cdot \bar{M}_{n-1} \neq 0$ and $\bar{C}_1 = \bar{M}_1 = \bar{\bar{C}}$, it is clear that the conditions (1) and (2) on the set M are satisfied since no $Q_{n_1,\ldots,n_k}$ belongs to an endelement of $\bar{\bar{C}}$ and $\bar{K}_{n_1,\ldots,n_k}$ has no points in common with $\bar{M} - K_{n_1,\ldots,n_k}$ except a single point of $\bar{\bar{C}}_{n_1,\ldots,n_{k-1}}$. We shall show that, for every $\epsilon > 0$, only a finite number of all the $\bar{K}_{n_1}, \bar{K}_{n_1,n_2}, \cdots, \bar{K}_{n_1,\ldots,n_k}, \cdots$ are of diameter $\geqq \epsilon$ and this will demonstrate (3) since each $\bar{\bar{C}}_{n_1,\ldots,n_k} \subset \bar{K}_{n_1,\ldots,n_k}$ and each component of $\bar{M} - \bar{M}_n$ is a $K_{n_1,\ldots,n_k}$ for some k. Then it will be clear (using (3)) that every point of $\bar{M} - M$ is a limit point of M, and hence, by W., Theorem 11, is an end point of $\bar{M}$. By P. S. T., chapter II, Theorem 42, this set is completely disconnected.

Now suppose there were an infinite number of the $\bar{K}$'s which were of diameter $\geq \epsilon > 0$. By repeated use of the preceding lemma we know that, for each k, only a finite number of the $\bar{K}_{n_1,\ldots,n_j}$, $1 \leq j \leq k$, are of diameter $\geq \epsilon$. Hence it is easy to see that we can find a sequence $\bar{K}_{\bar{n}_1} \supset K_{\bar{n}_1,\bar{n}_2} \supset \cdots$, all of which are of diameter $\geq \epsilon$, each $\bar{K}_{\bar{n}_1,\ldots,\bar{n}_k}$ being a particular one of the $\bar{K}_{n_1,\ldots,n_k}$. Let $\bar{K} = \prod_{k=1}^{\infty} \bar{K}_{\bar{n}_1,\ldots,\bar{n}_k}$; $\bar{K}$ is of diameter $\geq \epsilon$. Now each $\bar{C}_{\bar{n}_1,\ldots,\bar{n}_k}$ is so chosen that it contains a point $P_{\bar{n}_1,\ldots,\bar{n}_k}$ as far as possible and hence at a distance $\geq \epsilon/2$ from $Q_{\bar{n}_1,\ldots,\bar{n}_k} = \bar{C}_{\bar{n}_1,\ldots,\bar{n}_k} \cdot \bar{C}_{\bar{n}_1,\ldots,\bar{n}_{k-1}}$. We have seen that any arc joining a point of $\bar{K}_{\bar{n}_1,\ldots,\bar{n}_k}$ to a point of $\bar{C}_{\bar{n}_1,\ldots,\bar{n}_{k-1}}$ must contain $Q_{\bar{n}_1,\ldots,\bar{n}_k}$. Hence it is easy to see that any arc joining a point of $\bar{K}$ to a point of $\bar{C}_{\bar{n}_1,\ldots,\bar{n}_k}$ must contain all the $Q_{\bar{n}_1,\ldots,\bar{n}_l}$, $l > k$. Then let $\{l_k\}$ be a subsequence of the integers such that

$$\lim_{k\to\infty} P_{\bar{n}_1,\ldots,\bar{n}_{l_k}} = P_0 \quad \text{and} \lim_{k\to\infty} Q_{\bar{n}_1,\ldots,\bar{n}_{l_k}} = Q_0.$$

Obviously $P_0 \in \bar{K}$, $Q_0 \in \bar{K}$, and $d(P_0, Q_0) \geqq \epsilon/2$. Since every arc joining $P_{\bar{n}_1,\ldots,\bar{n}_{l_k}}$ to P_0 must go through all the points $Q_{\bar{n}_1,\ldots,\bar{n}_{l_m}}$, $m > k$, it must go through Q_0. But this contradicts Lemma 2.

2. *Hemicactoids and correspondences.* In this section, we shall intro-

duce the notion of a "hemicactoid" and bring out the relationship between hemicactoids and "upper semicontinuous collections of mutually exclusive continua filling up" a plane Jordan region. The connection between the theory of these collections and the theory of surfaces is suggested by Theorem 1, below.

Definition 1 (P. S. T., p. 28). We shall say that a continuum $\bar{C}$, is *the limit continuum* of the sequence $\{\bar{C}_n\}$ of continua if

(i) all the limit points of a sequence $\{P_n\}$ of points, $P_n \varepsilon \bar{C}_n$, lie in $\bar{C}$, and

(ii) if P is any point of $\bar{C}$, there exists a sequence $\{P_k\}$, $P_k \varepsilon \bar{C}_{n_k}$, of points which converges to P, $\{n_k\}$ being a subsequence of the integers.

LEMMA 1 (P. S. T., chapter I, Theorem 42). *If $\{\bar{C}_n\}$ is a sequence of continua such that there exists a convergent sequence $\{P_n\}$ of points, $P_n \varepsilon \bar{C}_n$, then the sequence $\{\bar{C}_n\}$ possesses a unique limit continuum $\bar{C}$, which consists of all points P satisfying* (ii), Definition 1.

Definition 2 (cf. P. S. T., chapter V, p. 324). A collection G of continua is *upper semicontinuous* provided it is true that if g is a continuum of G and $\{g_n\}$ is a sequence of continua of G, each g_n containing a point p_n such that the sequence $\{p_n\}$ converges to a point p of g, then g contains the limit continuum of the g_n.

Definition 3. Given an upper semicontinuous collection G of continua, a continuum g of G, and a sequence $\{g_n\}$ of continua of G. g will be said to be the *unique limit continuum with respect to* G of the sequence $\{g_n\}$ if and only if every limit point of any sequence $\{p_n\}$ of points where $p_n \varepsilon g_n$ for each n, lies in g. If there is a sequence $\{p_n\}$ of points, where $p_n \varepsilon g_n$ which converges to a point p of g, it is clear from Definition 2 that g will be the *unique* limit element of $\{g_n\}$ (with respect to G).

Remarks. Using P. S. T., chapter V, Theorems 1 and 2, we see that this definition is equivalent to that given in P. S. T., page 326. Clearly, if the continua of a collection are all points, this notion of limit element reduces to the ordinary notion of limit point. It is also clear that the notions of closed set and connected set of elements of G may be defined in terms of limit element.

LEMMA 2 (P. S. T., chapter V, Theorems 4 and 5). *A necessary and sufficient condition that a set H, of elements of an upper semicontinuous collection of mutually exclusive continua be closed or connected is that the set of points covered by H be closed or connected, respectively.*

Definition 4. A correspondence between the elements of two upper semicontinuous collections of mutually exclusive continua will be said to be *topological* if it is $1-1$ and limit elements are preserved.

Remark. It is clear that if T_1 is a topological transformation of G_0 into G_1 and T_2 of G_1 into G_2, then $T_1 \cdot T_2$ is a topological transformation of G_0 into G_2 if the G_i are all upper semicontinuous collections of mutually exclusive continua (which may all reduce to points in a given collection).

Definition 5. A vector function, $U = \phi(u)$, defined on a continuum $\bar{C}$, is said to be *monotone* if it is continuous and for every U_0, the set c_0 of points u for which $\phi(u) = U_0$ is null or a continuum.

THEOREM 1. *If $\phi(u)$ is a continuous vector function defined over a connected set C which is such that the sets c_0 of points u of C for which $\phi(u) = \phi(u_0)$ are all closed, then the maximal continua over each of which $\phi(u)$ is constant form an upper semicontinuous collection of mutually exclusive continua filling up C.*

Proof. Let $\{c_n\}$ be a sequence of the above continua such that there exists a sequence $\{p_n\}$ of points, p_n being in c_n for each n, which converges to a point p of C and let c be the continuum of the collection which contains p. Let $\bar{c}$ be the limit continuum of $\{c_n\}$ and let $q \in \bar{c}$; $p \in \bar{c}$ of course. There exists a sequence $\{q_{n_k}\}$, $q_{n_k} \in c_{n_k}$ converging to q. Then

$$\phi(q) = \lim_{k\to\infty} \phi(q_{n_k}) = \lim_{k\to\infty} \phi(p_{n_k}) = \phi(p),$$

and thus $\phi(u)$ is constant over $\bar{c}$. Hence $\bar{c} \subset c$ and thus the given collection is upper semicontinuous.

THEOREM 2. *Let G be any upper semicontinuous collection of continua filling up a connected set C which is the outer limit of a sequence of continua and let T be a transformation of the elements of G into the points of a set K. Let $U(u)$ be the vector function defined on C by the condition that $U(u) = U_P$ when u belongs to a continuum of G which is carried by T into the point P. Then a necessary and sufficient condition that T be topological is* (1) *that $U(u)$ be monotone and* (2) *the collection of maximal continua over each of which $U(u)$ is constant is identical with G.*

Proof. Suppose T is topological. Let $\{u_n\}$ be a sequence of points of C converging to the point u_0 of C, let c_0 be the continuum of G containing u_0, and let c_n be that containing u_n, for each n. From the definition of limit element of G, it follows that c_0 is the unique limit element of $\{c_n\}$. Hence $U(u_0) = \lim_{n\to\infty} U(u_n)$ and $U(u)$ is thus continuous. Furthermore, since T is

$1-1$, every point u of C for which $U(u) = U(u_0)$ lies in c_0. Hence $U(u)$ is monotone and (2) is satisfied.

Now suppose conditions (1) and (2) are satisfied. Let C be the outer limit of the continua $\bar{C}'_m$, let $\bar{C}'_m \subset \bar{C}'_{m+1}$ for each m, and, for each m, let G_m be the subcollection of G consisting of those continua of G which have a point in common with $\bar{C}'_m$. Clearly the set $\bar{C}_m$ of points covered by G_m is a continuum since it is obviously connected and is closed by P. S. T., chapter V, Theorem 2. Furthermore G_m is upper semicontinuous by P. S. T., chapter V, Theorem 1. Also (1) C is the outer limit of the continua $\bar{C}_m$, (2) the set of points $\bar{K}_m$ into which $\bar{C}_m$ is carried by $U = U(u)$ is a continuum for each m and (3) K is the outer limit of the $\bar{K}_m$. Hence if we show that T, $T : U = U(u)$, yields a topological transformation of G_m into $\bar{K}_m$ for each m, the desired result will follow.

Clearly T is the transformation of the maximal continua of $\bar{C}_m$ over each of which $U(u)$ is constant into the points of $\bar{K}_m$ and is therefore $1-1$, $U(u)$ being monotone. The continuity of T from G_m to $\bar{K}_m$ is immediate. Now let $\{P_n\}$ be a sequence of points of $\bar{K}_m$ converging to the point P_0 (of $\bar{K}_m$), let c_0 be the continuum of G_m corresponding to P_0, and for each n let c_n be that corresponding to P_n. Let $\{p_{n_k}\}$ be a convergent sequence of points, $p_{n_k} \in c_{n_k}$. The limit point p_0 of this sequence is in $\bar{C}_m$ and, since $U(u)$ is continuous, $U(u_{p_0}) = \lim_{k\to\infty} U(u_{p_{n_k}}) = U_{p_0}$ so that $p_0 \in c_0$. Thus c_0 is the limit continuum of $\{c_n\}$ in G_m and hence T is continuous.

LEMMA 3. *A necessary and sufficient condition that a plane continuum should fail to separate the plane is that it be representable as the product of a sequence $\{\bar{R}_n\}$ of closed Jordan regions each of which contains the next.*

Proof. The condition is obviously sufficient.

Now suppose that the continuum $\bar{C}$ does not separate the plane Π. Then any point P of $\Pi - \bar{C}$ can be joined to ∞ by an arc, all of whose points are at a distance $\geq d_P > 0$ from $\bar{C}$. Let $\bar{R}_n$ be the set of all points bounded by the outer simple closed curve of the set $\bar{R}'_n$ of points at a distance $\leq 1/n$ from $\bar{C}$; $\bar{R}'_n$ is known to be a region bounded by a finite number of simple closed curves. Then $\bar{C} = \prod_{n=1}^{\infty} \bar{R}_n$, for any point P in $\bar{C}$ is in each $\bar{R}_n$ and any point P in $\Pi - C$ is outside of $\bar{R}_n$ for $n > 1/d_P$.

Definition 6. By a *base set* we shall mean a bounded continuous curve in the plane which does not separate the plane.

LEMMA 4. *Let G be an upper semicontinuous collection of mutually exclusive continua filling up a Jordan region $\bar{r}$, no one of which separates the*

plane although some may separate $\bar{r}$. Then G is topologically equivalent to a base set.

Proof. Let G_1 be the collection, filling up the plane Π, consisting of G plus all the points of Π outside $\bar{r}$. Clearly G_1 is upper semicontinuous and no continuum of G_1 separates Π. From a theorem of R. L. Moore,† we can find a $1-1$ continuous mapping $T, T : U = U(u)$, $U(u)$ monotone, of the continua of G_1 on to the points of Π. Let $\{r^*_n\}$ be a sequence of simple closed curves approaching r^* uniformly, each containing the next and r^* in its interior. Now evidently r^*_n is carried in a $1-1$ continuous way into a Jordan curve R^*_n for each n. Now no point interior to r^*_n corresponds to a point exterior to R^*_n for first, we can find points interior and near to r^*_n which are carried interior to R^*_n and second, the set of points corresponding to r_n is connected. Similarly, points outside r^*_n go into points outside R^*_n and thus T carries all of r_n into all of R_n and all of $\Pi - \bar{r}_n$ into all of $\Pi - \bar{R}_n$, $\bar{R}_n$ being the Jordan region bounded by R^*_n. Hence each $\bar{R}_n$ contains the next, all contain $T(\bar{r})$, and any point not in $T(\bar{r})$ is outside of $\bar{R}_n$ for some n so that $T(\bar{r}) = \prod_{n=1}^{\infty} \bar{R}_n$. That $T(\bar{r})$ is a continuous curve follows from the fact that it is the range of values of the continuous vector function $U(u)$ (see Lemma 3, § 1).

Definition 7. By a *double cone*, we shall mean the surface of revolution generated by revolving an isosceles triangle about its base as axis. By the *axis* of the double cone, we mean its axis of revolution; this segment is not part of the surface. By the *end points* of the double cone, we mean the end points of its axis, and by its *vertex angle*, the magnitude of a base angle of a generating isosceles triangle.

Definition 8. Given a rhombus, we shall designate two opposite vertices as "*end points*," define its *vertex angle* as half the angle of the rhombus subtended at one of the end points, and define its *axis* as the diagonal segment joining its end points.

Definition 9. By a *simple cyclic chain of type A* (*type B*) we shall mean a point set obtained as follows: let L be a closed line segment and F a closed set on it; replace each complementary interval of F on L by a double cone (rhombus) with the same end points, the vertex angles of the double cones (rhombuses) being the same.

Definition 10 (P. S. T., p. 151). By a *cactoid*, we mean a bounded con-

† R. L. Moore, "Concerning upper semicontinuous collections of continua," *Transactions of the American Mathematical Society*, vol. 27 (1925), p. 425.

tinuous curve (in a metric space) all of whose non-degenerate simple links are simple closed surfaces.

Definition 11. By a *canonical cactoid* (*base set*), we mean a bounded continuous curve in 3-space which may be expressed in the form in Theorem 1, § 1, the $\bar{C}_n$ being all simple cyclic chains of type A (type B).

LEMMA 5. *Any simple cyclic chain of a cactoid* (*base set*) *can be mapped in a* $1-1$ *continuous way on a simple cyclic chain of type A* (*type B*).

Proof. From the definition of cactoid and from the fact (Lemma 4, § 1) that every non-degenerate simple link of a continuous curve $\bar{M}$ is a non-degenerate cyclic element of $\bar{M}$ and conversely, we see that every maximal cyclic curve of a cactoid is a simple closed surface. From W., Theorem 24, we see that a necessary and sufficient condition that a plane continuous curve fail to separate the plane is that each of its maximal cyclic curves is a closed Jordan region.

Now let $\bar{C}$ be the simple cyclic chain under consideration and let E_1 and E_2 be its endelements. Let $P_i = E_i$ if E_i is a point or let P_i be an internal point (Definition 12, § 1) of E_i if E_i is non-degenerate, $(i = 1, 2)$, and let t be an arc joining P_1 and P_2. Let t be mapped in a $1-1$ continuous way on a line segment L of the desired length. Let K be the set of points separating P_1 from P_2 in $\bar{C}$; by Lemma 7, § 1, all the points of K lie on t and $K + P_1 + P_2$ is closed. Let Σ be the closed set on L corresponding to $K + P_1 + P_2$. By Lemma 7, § 1 and W., Theorem 1, each arc of t corresponding to a complementary interval of Σ on L determines a unique maximal cyclic curve of $\bar{C}$ and $\bar{C}$ is the sum of $K + E_1 + E_2$ and all of these maximal cyclic curves. Furthermore, if an E_i is non-degenerate, there is an arc of t, with P_i as end point, lying in E_i in which case, we say that E_i is determined by this arc. Now map each of these maximal cyclic curves of $\bar{C}$ corresponding to a given arc of t on the double cone (if $\bar{C}$ belongs to a cactoid) or rhombus (if $\bar{C}$ belongs to a base set) which has the corresponding end points on L and the given vertex angle. Clearly this $1-1$ map of $\bar{C}$ on $\bar{\Gamma}$, the canonical chain, is continuous for, by the argument of Lemma 8, § 1, there are only a finite number of maximal cyclic curves of $\bar{C}$ having a diameter $\geq \epsilon$, for each $\epsilon > 0$.

Definition 12. A *hemicactoid* is a continuous curve in 3-space which is the sum of a canonical base set $\bar{B}$, in the (x, y) plane Π and a finite or denumerable set of canonical cactoids $\bar{K}_n$, each one of which has exactly one point in common with $\bar{B}$, is otherwise entirely above Π, and has no point, not in $\bar{B}$, in common with any of the other cactoids.

Definition 13. Two hemicactoids are said to be *homeomorphic* if and

only if they can be put into a $1-1$ continuous correspondence in which the base sets are also in a $1-1$ continuous correspondence.

The following lemma is immediate:

LEMMA 6. *Let* $\bar{M}$ *be a hemicactoid* $\bar{M} = \bar{B} + \sum_{m=1}^{(\infty)} \bar{K}_m + M'$, *(the symbol* (∞) *means that the upper limit may be finite or infinite) the number of simple cyclic chains in either* $\bar{B}$ *or the totality of the* $\bar{K}_m$ *being finite. Let* (1) $\epsilon > 0$, (2) $\eta > 0$, (3) $\bar{\Gamma}$ *be a simple cyclic chain of type A if the totality of the* $\bar{K}_m$ *contains only a finite number of chains or of type B if this is not the case, and* (4) P *be a point of one of the* $\bar{K}_m$ *not on an end cyclic element other than* $\bar{K}_m \cdot \bar{B}$ *of a* $\bar{K}_m$ *in the first case above, or any point of* B^* *not on an end cyclic element of* $\bar{B}_1$ *(one chain of* $\bar{B}$*) in the second case. Then we can find a canonical (i. e. of type A or B) simple cyclic chain* $\bar{C}$, *homeomorphic to* $\bar{\Gamma}$, *which may be attached to* $\bar{M}$ *at* P *in such a way that* $\bar{M} + \bar{C}$ *is a hemicactoid and, for* $\rho \geq \epsilon$, $\eta_{\bar{M}+\bar{C}}(\rho) > (1-\eta)\eta_{\bar{\Lambda}}(\rho)$, *and for* $0 < \rho < \epsilon$, $\eta_{\bar{M}+\bar{C}}(\rho) > 0$, *where* $\eta_{\bar{L}}(\rho)$ *is the lower boundary, for* $d_{\bar{L}}(P, Q) \geq \rho$ *of the ratio of* $d(P, Q)$ *to* $d_{\bar{L}}(P, Q)$, $d_{\bar{L}}(P, Q)$ *being the geodesic distance along* $\bar{L}$ *from* P *to* Q.

LEMMA 7. *Any cactoid (base set) can be mapped in a* $1-1$ *continuous way on a canonical cactoid (base set).*

Proof. Let $\bar{M}'$ be the given cactoid (base set) and let it be represented as in Theorem 1, § 1. By Lemma 5, each $\bar{C}'_n$ can be mapped in a $1-1$ continuous way on a simple cyclic chain of type A (type B). Let $\bar{C}'_1$ be so mapped on such a chain $\bar{C}_1$ of unit length and vertex angle $\leq \pi/4$. By Lemmas 5 and 6, we can find a simple cyclic chain $\bar{C}_2$, homeomorphic to $\bar{C}'_2$, of length $\leq 1/4$ and vertex angle $\leq \pi/4$ which can be attached to $\bar{C}_1$ at the point of $\bar{C}_1$ (not on an endelement of $\bar{C}_1$) corresponding to the point $\bar{C}'_1 \cdot \bar{C}'_2$ in such a way that $\bar{C}_1 + \bar{C}_2$ is a canonical cactoid (base set) and $\eta_2(\rho) = \eta_{\bar{C}_1+\bar{C}_2}(\rho) \geq (1 - 1/2^2)\eta_1(\rho)$ for $\rho \geq 1/2$, and $\eta_2(\rho) > 0$ for $0 < \rho < 1/2$, $\eta_1(\rho)$ being $\eta_{\bar{C}_1}(\rho)$. It is clear from Lemma 6 that this process may be continued indefinitely with $\epsilon = 1/n$, $\eta = 1/n^2$, the length of $\bar{C}_n \leq 1/n^2$, and the vertex angle of $\bar{C}_n \leq \pi/4$. At each stage, the correspondence between $\bar{M}'_n$ and M_n set up by uniting these individual correspondences between the chains is $1-1$ and uniformly continuous both ways. Also these unite to give a $1-1$ continuous correspondence between M and M' (cf. Theorem 1, § 1). Furthermore it is clear that there exists a continuous monotone function $\eta(\rho)$, positive for $\rho > 0$ and vanishing with ρ, such that $\eta(\rho) \leq \eta_n(\rho)$ for each n. We shall show † that this correspondence is uniformly continuous both ways

† The argument from this point on is essentially that used by H. M. Gehman in his

so that it can be extended to the closures, $\bar{M}$ and $\bar{M}'$ of M and M' respectively which shows also that $\bar{M}$ is a continuous curve and thus a canonical cactoid (base set).

Suppose the mapping of M onto M' is not uniformly continuous. Then there exist sequences $\{P_i\}$ and $\{Q_i\}$ of points of M and an $\epsilon > 0$ such that $d(P_i, Q_i) < 1/i$ and $d(P'_i, Q'_i) \geq \epsilon$. Let n be so large that the diameter of any component of $M' - \bar{M}'_n < \epsilon/3$. Then any arc of M' joining P'_i and Q'_i contains an arc $\widehat{E'_iF'_i}$ in $\bar{M}_n$ (entirely, by W., Theorem 8) of diameter $> \epsilon/3$. On the other hand there exists a $\delta > 0$ such that if $d(A, B) < \delta$, $A, B \,\varepsilon\, \bar{M}_n$, then $d(A', B') < \epsilon/3$. Also, since M has the property that any two of its points at a distance $\leq \rho$ apart can be joined by an arc of M of length $\leq \eta^{-1}(\rho)$, $\eta^{-1}(\rho)$ being the inverse function of $\eta(\rho)$ and approaching zero with ρ, we see by W., Theorem 9, that M is uniformly connected in kleinen and there exists an $\alpha > 0$ such that if $d(A, B) < \alpha$, $A, B \,\varepsilon\, M$, A and B can be joined by an arc of M of diameter $< \delta$. Hence if $i > 1/\alpha$, $d(P_i, Q_i) < \alpha$ and they can be joined in M by an arc of diameter $< \delta$ so that the arc $\widehat{E_iF_i}$ in $\bar{M}_n$ is of diameter $< \delta$ and thus the diameter of $\widehat{E'_i, F'_i} < \epsilon/3$, contrary to the third sentence of this paragraph.

Now, it is clear that, for every $\epsilon > 0$, there exists an $N(\epsilon)$ such that, for $n > N(\epsilon)$, the diameter of each maximal connected subset of $M - \bar{M}_n < \epsilon/3$. Also, by Lemma 2, § 1, $\bar{M}'$ is uniformly arcwise connected in kleinen. Hence the above argument can be used to demonstrate the uniform continuity of the mapping of M' on M. This is what we wished to prove.

LEMMA 8. *Let G be an upper semicontinuous collection of mutually exclusive continua filling up $\bar{r}$ and let G_0 be the subcollection of G consisting of the continua g_0 of G such that no point of g_0 is in a bounded complementary domain † of any continuum of G. Then every point P of $\bar{r}$ is either in a g_0 or in a bounded complementary domain of some g_0.*

Proof. Let G be the given collection and G_0 the above subcollection. It is clear that if one point of a continuum g of G is in a complementary domain of another continuum of G, all of g is in that complementary domain. Now a given point P of $\bar{r}$ is either in a g_0 or is in a bounded complementary domain of some continuum $g^{(1)}$ of G; either $g^{(1)} \,\varepsilon\, G_0$ or it lies wholly in a

---

paper "Concerning acyclic continuous curves," *Transactions of the American Mathematical Society*, vol. 29 (1927), pp. 553-555 in particular.

† A *complementary domain* of a continuum $\bar{C}$ in the plane Π is a component of $\Pi - C$. A bounded complementary domain of a continuum is an open simply connected region.

complementary domain of a continuum $g^{(2)}$ of G which also contains a circular disc outside of $g^{(1)}$. By a repetition of this argument, we obtain a normally ordered set $g^{(1)}, g^{(2)}, \cdots, g^{(\omega)} g^{(\omega+1)}, \cdots$, each $g^{(\alpha)} \in G$ and being wholly inside a bounded complementary domain of $g^{(\alpha+1)}$ which also contains a circular disc outside of $g^{(\alpha)}$. Thus this set is denumerable and hence there is a first ordinal β in Cantor's second class for which $g^{(\beta)} \in G_0$. This proves the lemma.

THEOREM 3. *Any upper semicontinuous collection of mutually exclusive continua filling up a Jordan region $\bar{r}$ is topologically equivalent to a hemicactoid. We may further require that a monotone transformation $U = U(u)$ of $\bar{r}$ into $\bar{H}$ carry the subcollection G_0 of Lemma 8 into the base set $\bar{B}$. If $U = U(u)$ and $U = U_1(u)$ are two such transformations of $\bar{r}$ into $\bar{H}$ and $\bar{H}_1$ respectively, then there exists a $1-1$ continuous transformation $U_1 = \bar{U}_1(U)$ of $\bar{H}$ into $\bar{H}_1$ which carries $\bar{B}$ into $\bar{B}_1$ topologically and is such that $\bar{U}_1[U(u)] = U_1(u)$.*

Proof. Now let G be the given collection and G_0 the subcollection of Lemma 8. If we form the collection G'_0 of continua g'_0, where each g'_0 is a g_0 plus all of its bounded complementary domains, we see that no continuum of G'_0 separates Π, G'_0 is upper semicontinuous and topologically equivalent to G_0 and, by Lemma 8, fills up $\bar{r}$. Thus G_0 is topologically equivalent to a canonical base set $\bar{B}$. Let $U = U(u)$ yield such a correspondence.

Now the subcollection of G which lies in a bounded complementary domain d of some g_0 plus that g_0 is topologically equivalent to an upper semicontinuous collection filling up the surface of a sphere, for there exists a monotone transformation of $\bar{d}$ into the sphere which carries d^* into a point and is otherwise $1-1$. Such a collection is known, from a theorem of R. L. Moore (P. S. T., Theorem 20, chapter VII) and from Lemma 7, to be topologically equivalent to a canonical cactoid. It is clear that there are at most a denumerable infinity of such bounded complementary domains, so let them be ordered into a single sequence $\{d_n\}$ and let $\bar{K}'_n$ be a canonical cactoid corresponding to $d_n + g_0^{(n)}$ ($g_0^{(n)}$ being the g_0 containing d^*_n), the correspondence being given by $U = U'_n(u)$. Let $\bar{K}'_n = \sum_{p=1}^{\infty} \bar{C}'_{n,p} + K'^*_n$, and let the $\bar{C}'_{n,p}$ be ordered into a single sequence $\{\bar{\bar{C}}_m\}$ such that

$$\bar{\bar{C}}_{m+1} \cdot (\bar{B} + \bar{\bar{C}}_1 + \cdots + \bar{\bar{C}}_m) \neq 0$$

for any $m \geq 0$. Now by Lemma 6, we can find a simple cyclic chain $\bar{C}_1$, of type A, homeomorphic to $\bar{\bar{C}}_1$ and of length ≤ 1, and can join it onto $\bar{B}$, projecting upwards, at the point of $\bar{B}$ corresponding to the g_0 containing the

boundary of the $\bar{d}_a$ in $\bar{r}$ corresponding to the cactoid containing $\bar{\bar{C}}_1$. Clearly $\bar{B} + \bar{\bar{C}}_1$ is a hemicactoid. Let $U(u)$ be extended to the part $\bar{d}_{1a}$ of $\bar{d}_a$ which corresponds to $\bar{\bar{C}}_1$ so that $U = U(u)$ gives a topological correspondence between $G_0 + G \cdot \bar{d}_{1a}$ and $\bar{B} + \bar{C}_1$. By Lemma 6, we can find a $\bar{C}_2$ homeomorphic to $\bar{\bar{C}}_2$ of length $\leq 1/4$, which can be attached to $\bar{B} + \bar{C}_1$ at the point corresponding to the continuum of $G_0 + G \cdot \bar{d}_{1a}$ which bounds the complementary domain of $\hat{G}_0 + \bar{d}_{1a}$ ($\hat{G}_0$ being the set of points covered by G_0) in which the continua corresponding to $\bar{\bar{C}}_2$ lie,† and is such that

$$\eta_{\bar{B}+\bar{C}_1+\bar{C}_2}(\rho) = \eta_2(\rho) \geq (1 - 1/2^2)\eta_1(\rho), \qquad \rho \geq 1/2,$$
$$\eta_2(\rho) > 0, \qquad 0 < \rho < 1/2,$$

where $\eta_1(\rho) = \eta_{\bar{B}+\bar{C}_1}(\rho)$. From Lemma 6, it is clear that this may be continued indefinitely so that the length of $\bar{C}_n \leq 1/n^2$ and $\eta_n(\rho) \geq [1 - 1/n^2]\eta_{n-1}(\rho)$ for $\rho \geq 1/n$ and $\eta_n(\rho) > 0$ for $0 < \rho < 1/n$. By the method of Lemma 7 it can be shown that the $\bar{C}_i$ corresponding to the $\bar{\bar{C}}_i$ of one of the $\bar{K}'_n$ plus the limit points of this set not in any $\bar{C}_i$ add up to a canonical cactoid $\bar{K}_n$ homeoiphic to $\bar{K}'_n$, that the closure $\bar{H}$ of the set $H = \bar{B} + \sum_{i=1}^{\infty} \bar{C}_i$ is a hemicactoid, that $\bar{H} = \bar{B} + \sum_{n=1}^{\infty} \bar{K}_n$, and that $U(u)$ extended as indicated to all of $\bar{r}$ remains continuous and yields a topological transformation of G into $\bar{H}$ and G_0 into $\bar{B}$, being monotone consequently.

The remaining statements are immediate, remembering the fact (remark after definition 4) that the product of two topological transformations of upper semicontinuous collections of continua is again topological and that the notions of limit point and limit element in a collection of continua are equivalent if the continua are all points (remarks after definition 3).

LEMMA 9. *Given a base set (cactoid) in n-space. There exists an upper semicontinuous collection of mutually exclusive continua filling up a Jordan region $\bar{r}$ (the surface Σ of a sphere) which is topologically equivalent to the given point set. In the case of the base set it may be further required that no continuum of the corresponding collection separates Π and also that every non-degenerate continuum separates $\bar{r}$. In the case of the cactoid, it may be assumed that each non-degenerate continuum of the collection separates Σ.*

Proof. The lemma for cactoids has already been proved by R. L. Moore

† That all of the continua corresponding to $\bar{\bar{C}}_2$ lie in such a bounded complementary domain follows from the arguments of Lemma 8.

(see P. S. T., chapter VII, Theorem 23). We shall give a brief proof of this part of the lemma using the result for base sets, which we shall now prove.

Let $\bar{B}$ ($\bar{B} = B' + \sum_{i=1}^{\infty} \bar{C}_i$, where the $\bar{C}_i$ are simple cyclic chains of type B, $\bar{C}_i \cdot \bar{C}_j$ is at most one point, $i \neq j$, and B' is completely disconnected as in Theorem 1, § 1) be a canonical base set in the plane Π, homeomorphic to the given base set. Let $\Pi - \bar{B}$ be mapped conformally on $\Pi - \bar{R}$, $\bar{R}$ being the the closed unit circle (clearly there is no loss of generality in taking $\bar{r} = \bar{R}$), by the vector function $U = U(u)$, $u \in \Pi - \bar{R}$. Now, by Lemma 1, § 3, $U(u)$ remains continuous on R^* and is not constant over any continuum of R^*. We shall proceed to extend $U(u)$ to all of $\bar{R}$ so that it is monotone, and $U = U(u)$ carries $\bar{R}$ into $\bar{B}$ in such a way that any non-degenerate continuum over which $U(u)$ is constant separates $\bar{R}$ (but not Π) and is carried by $U = U(u)$ into a cut point of $\bar{B}$.

Let Σ_1 be the closed sub-set of R^* which is carried into $(\Pi - \bar{C}_1)^*$. It is clear from the well known nature of $U(u)$ on $(\Pi - \bar{R})^*$, that the end points of each complementary interval of Σ_1 correspond to the same point of $(\Pi - \bar{C}_1)^*$. Hence join each of these pairs of points by a segment and define $U(u)$ to be constant over each such segment; let R_1 be the convex region bounded by the simple closed curve formed by these segments and the points of R^* not cut off by any of them. If we now define the elements of R^*_1 to be the continua of R^*_1 over which $U(u)$ is constant, it is seen that at most two of these elements are carried into the same point of $(\Pi - \bar{C}_1)^*$. Let each point of each element of a corresponding pair be joined to each point of the corresponding element by a segment, and define $U(u)$ to be constant over each such segment. Now, the remainder of R_1 consists of Jordan regions ρ_1 such that $U = U(u)$ gives a representation of the boundary of a rhombus of $\bar{C}_1$ on the boundary ρ^*_1; hence define $U(u)$ to be continuous in each such $\bar{\rho}_1$ and such that $U = U(u)$ gives a $1-1$ continuous map of the interior of the rhombus on the interior of ρ_1.

Now, a unique one of the segments on the boundary of R_1 is carried by $U = U(u)$, so far defined, into the unique point $\bar{C}_1 \cdot \bar{C}_2$ and has the additional property that it cuts off all the points of R^* carried into points of $\bar{C}_2 - \bar{C}_1 \cdot \bar{C}_2$. Let Σ_2 be this segment plus the set of all these points of R^*. Again, each pair of end points of a complementary interval of Σ_2 corresponds to the same point of $\bar{C}_2$. Join all these pairs of points by line segments and define $U(u)$ to be constant over each such segment. Let R_2 be the convex region bounded by all these new segments and Σ_2. If each continuum of R^*_2 over which $U(u)$ is constant be considered as an element, it is again seen that at most two such elements are carried by $U = U(u)$ into the same point of $\bar{C}_2$. Join each point

of each element of a pair of corresponding elements to each point of the corresponding element by a segment and define $U(u)$ to be constant over each such segment. Clearly the rhombuses of $\bar{C}_2$ may be mapped in a $1-1$ continuous way on the remaining regions of R_2, as before.

It is now clear that this may be continued indefinitely and that finally $U(u)$ will be defined over all of $\bar{R}$. It is clear that it will be continuous, since, for every $\epsilon > 0$, at most a finite number of the $\bar{C}_n$ and a finite number of the rhombuses are of diameter $\geq \epsilon$. It is also immediate that it is monotone and that every continuum of $\bar{R}$ which is carried into a point of $\bar{B}$ separates $\bar{R}$ and is carried into a cut point of $\bar{B}$.

To prove the lemma for cactoids, we first replace the cactoid by a homeomorphic canonical one $\bar{C}$, the axes of whose simple cyclic chains form an acyclic continuous curve lying entirely in Π (that this can be done follows from Lemmas 5 and 6 and the argument of Lemma 7). Now, we map the closed upper half of $\bar{C}$ on the closed upper half of a hemisphere by projecting it onto Π, forming a canonical base set $\bar{B}$, mapping $\bar{B}$ on $\bar{R}$, and then projecting up onto the hemisphere from $\bar{R}$. We then extend the mapping by making a point P on the lower half of the cactoid to correspond to that point on the lower hemisphere which is the reflection in Π of the point of the upper hemisphere corresponding to the reflection of P on the cactoid.

THEOREM 4. *Given any hemicactoid $\bar{H}$, we can find a monotone vector function $U(u)$ such that $U = U(u)$ carries $\bar{r}$ into $\bar{H}$. If G is the upper semicontinuous collection of continua over each of which $U(u)$ is constant, i. e., the collection corresponding to $U(u)$, and G_0 is the subcollection of Lemma 8, we may require that $U = U(u)$ carry G_0 into $\bar{B}$.*

Proof. By Lemma 9, we can find an upper semicontinuous collection $G_0^{(0)}$ filling up $\bar{r}$ which is topologically equivalent to $\bar{B}$ and such that none of its continua separate the plane. Let $U = U_0(u)$ be a monotone transformation giving such a correspondence. Now order the cactoids of $\bar{H}$ having only one point each in common with $\bar{B}$ into a sequence $\{\bar{K}_n\}$. Let us take $\bar{r} = \bar{Q}$ (obviously we may take $\bar{r} = \bar{Q}$ without loss of generality). Now $\bar{K}_1$ joins on to $\bar{B}$ at a point P_1; let p_1 be a point of the continuum of $G_0^{(0)}$ corresponding to P_1. Make a transformation of $\bar{Q}$ into itself defined as follows: each point $p \neq p_1$ shall correspond to the point p' on the directed ray p_1p such that $d(p', p'') = d(p, p'')/2$, p'' being the last point of $\bar{Q}$ on the ray p_1p; the point p_1 will be carried into all the points p' on the boundary of the square $d(p_1, p') = d(p_1, p'')/2$, p'' being the last point of $\bar{Q}$ on the ray p_1p'. Now define $U_1(u)$ outside this square as the transform of $U_0(u)$ and define it inside and on this square so as to be continuous in $\bar{Q}$ and carry

the closed square in a monotone way into $\bar{K}_1$. That this can be done follows from the fact that we may carry it in a monotone way onto the surface of a sphere Σ, the boundary being carried into a point and then, by Lemma 9, we can carry Σ in a monotone way into $\bar{K}_1$. $U_1(u)$ is clearly monotone.

Now $\bar{C}_2$ is attached to $\bar{B}$ at a point P_2 (perhaps $= P_1$); let p_2 be a point corresponding under $U = U_1(u)$ to P_2. Make a transformation of $\bar{Q}$ into itself similar to the above in which any point $p \neq p_2$ is carried into the point p' on the ray p_2p such that $d(p', p'') = (1 - 1/2^2)d(p, p'')$ and p_2 is carried into all the points p' satisfying $d(p_2, p') = d(p_2, p'')/2^2$, p'' being the last point of $\bar{Q}$ on p_2p'. Define $U_2(u)$ outside this square as the transform of $U_1(u)$ and within and on the square so as to be continuous in $\bar{Q}$ and carry the closed square in a monotone way into $\bar{K}_2$. Clearly $U_2(u)$ is monotone and $U = U_2(u)$ carries $\bar{Q}$ into $\bar{B} + \bar{K}_1 + \bar{K}_2$ and the twice transformed collection $G_0^{(2)}$ topologically into $\bar{B}$.

Let this be continued indefinitely. A simple analysis serves to show that the sequence $\{U_n(u)\}$ converges uniformly to a monotone function $U(u)$ with corresponding collection G, that $G_0 = G_0^{(\infty)}$, and that $U = U(u)$ carries $\bar{Q}$ into $\bar{H}$ and G_0 topologically into $\bar{B}$.

THEOREM 5. *Let $\bar{H}$ be a hemicactoid with base set $\bar{B}$ and canonical cactoids $\bar{C}_m$; we have $\bar{B} = \sum_{n=1}^{\infty} \bar{B}_n + B'$ and $\bar{C}_m = \sum_{n=1}^{\infty} \bar{C}_{m,n} + C'_m$, where the B' and C'_m are mutually exclusive completely disconnected sets and $\bar{B}_n$ and $\bar{C}_{m,n}$ are simple cyclic chains of types B and A respectively joined together as in Definition 12 and Theorem 1, § 1. Now on each $\bar{B}_n$ and $\bar{C}_{m,n}$, let there be determined a continuous monotone transformation $T_n^{(2)}$ (for $\bar{B}_n$) and $T_{m,n}^{(1)}$ (for $\bar{C}_{m,n}$) of each into itself, which carries each non-degenerate cyclic element into itself and carries the end points into themselves; we do not assume that the transformation is $1-1$ necessarily. Then we can find a hemicactoid $\bar{H}'$ with base $\bar{B}'$ and cactoids $\bar{C}'_m$ such that (1) there are $1-1$ continuous transformations $S_{m,n}^{(1)}$ of $\bar{C}_{m,n}$ into $\bar{C}'_{m,n}$ and $S_n^{(2)}$ of $\bar{B}_n$ into $\bar{B}'_n$ which are conformal (with the obvious conventions at vertices) between corresponding non-degenerate cyclic elements and (2) the transformations $[S_{m,n}^{(1)}]^{-1} T_{m,n}^{(1)}$ and $[S_n^{(2)}]^{-1} T_n^{(2)}$, defined respectively on $\bar{C}'_{m,n}$ and $\bar{B}'_n$, unite to form a continuous monotone transformation of $\bar{H}'$ into $\bar{H}$ which carries $\bar{B}'$ into $\bar{B}$.*

Proof. Let $\bar{B}'_1 = \bar{B}_1$. By Lemma 6, we can attach a simple cyclic chain $\bar{B}'_2$ homeomorphic to $\bar{B}_2$ at a point of $\bar{B}'_1$ which is carried by $(S_1^{(2)})^{-1} T_1^{(2)}$ ($S_1^{(2)}$ being taken to be the identity) into the point $\bar{B}_1 \cdot \bar{B}_2$; clearly we may assume that the transformation $S_2^{(2)}$ of $\bar{B}_2$ into $\bar{B}'_2$ is conformal between corresponding non-degenerate cyclic elements. It is clear by Lemma 6 that

this can be continued indefinitely for the $\bar{B}_n$ and then for the $\bar{C}_{m,n}$ arranged in a single sequence. By arguments already used, all the conclusions may be demonstrated.

3. *Monotone transformations in the plane.* In this section, we shall merely develop a few important theorems about monotone transformations of one plane set into another.

LEMMA 1. *Let $w = f(z)$, $f(z)$ being analytic in R_0, the unit circle except for a possible single pole, carry R_0 in a $1-1$ conformal way into a simply connected region D which may be infinite. Then a necessary and sufficient condition that $f(z)$ be continuous on R^*_0 is that D^* be a bounded continuous curve. If this is the case, no continuum of R^*_0 corresponds to a point of D^*.*

Proof. That the condition is necessary is obvious.

Let us assume, then, that D^* is a bounded continuous curve. Then every point of D^* is accessible from D.† Thus each prime end ‡ of D^* consists of a single point. From the results of Caratheodory,‡ one immediately concludes that $f(z)$ is continuous over R^*_0 and that no continuum of R^*_0 corresponds to a point of D^*.

LEMMA 2. *Let $U = U(u)$ be a monotone transformation carrying a Jordan region $\bar{r}$, into a base set $\bar{B}$, in which no continuum over which $U(u)$ is constant separates Π. Then we can find a monotone transformation $U = U'(u)$, of the whole plane Π, into itself such that $U'(u) = U(u)$, $u \in \bar{r}$, and $U = U'(u)$ is $1-1$ and continuous for $u \in \Pi - \bar{r}$.*

Proof. If $\bar{B}$ consists of a single point, the theorem is immediate, for we may define G_1 to be the upper semicontinuous collection of continua consisting of $\bar{r}$ plus all the points of $\Pi - \bar{r}$. From a theorem of R. L. Moore,§ we can find a monotone transformation $U = U''(u)$ of Π into itself carrying G_1 topologically into the points of Π, and $\bar{r}$ into some point U_0''. If we define $U'(u) = U''(u) + \bar{U} - U_0''$, $(\bar{U} = U(u), u \in \bar{r})$, $U = U'(u)$ satisfies the requirements.

Now if $\bar{B}$ does not consist of a single point, map $\Pi - \bar{B}$ in a $1-1$ conformal way on $\Pi - \bar{r}$ by a function $U = U''(u')$; $U = U''(u')$ is continuous on r^* and no continuum of r^* corresponds to a point of $(\Pi - \bar{B})^*$, by Lemma 1. Now it is well known that there exists a monotone transformation of Π into itself, which is $1-1$ and continuous off of r^*, and

† R. L. Wilder, *loc. cit.*, pp. 350-351.

‡ C. Caratheodory, "Die Begrenzung einfache zusammenhängender Gebiete," *Mathematische Annalen*, vol. 73 (1913), pp. 323-351 and footnote, page 251, especially.

§ R. L. Moore, *loc. cit.* (not P. S. T.).

carries any upper semicontinuous collection of continua on r^*, none of which separates r^*, topologically into the points of r^*. Now, the collection of maximal continua of r^* each of which corresponds to a point of $(\Pi - \bar{B})^*$ is upper semicontinuous and no one includes all of r^* or separates r^*. Hence we have such a transformation $u' = u'(u)$, such that $U''(u') = U(u)$, $u \in r^*$. Hence if we define $U'(u) = U''[u'(u)]$, $u \in \Pi - \bar{r}$, and $U'(u) = U(u)$, $u \in \bar{r}$, then $U'(u)$ is continuous and monotone over all of Π and satisfies the other requirements.

Definition 1. By the oscillation of a vector function $\phi(u)$, over a set E, we mean the least upper bound of $|\phi(u_1) - \phi(u_2)|$ for all pairs (u_1, u_2) of points of E.

THEOREM 1. *Let $T : U = U(u)$, be a continuous monotone transformation of an open simply connected region r, into an open simply connected region R, and let G be the collection of maximal continua over each of which $U(u)$ is constant. Then*

(i) *no continuum of G separates r;*

(ii) *the points of any closed Jordan region $\bar{d}$, are carried by T into points of the continuum $\bar{D} = \tilde{D} + \tilde{D}^*$, where $\tilde{D}^* = T(d^*)$ and all points (if any) of $\tilde{D}$ are separated from ∞ by $\tilde{D}^*$;*

(iii) *if $\bar{d}$ is any closed region in r, the oscillation of $U(u)$ over $\bar{d}$ is equal to its oscillation over d^*; and*

(iv) *all the points of r carried into a simply connected open region D, of R, form a simply connected open region d, of r.*

If we replace r and R by closed Jordan regions $\bar{r}$ and $\bar{R}$, in the hypotheses, all the above conclusions hold, and we may add that no continuum of G contains all of r^*.

Proof. (i) is immediate since T is a topological transformation of G into R and since by Lemma 2, § 2, the set of points of r belonging to a connected set of elements of G is connected.

To demonstrate (ii), let P be a point of R which can be joined to a point Q of R^* by an arc Γ (open at one end) containing no point of $\tilde{D}^*$. To this arc corresponds a connected set γ, of points of r having no point in common with d^*. Let the arc Γ be the outer limit of arcs $\Gamma_1 + \Gamma_2 + \cdots$, where $\Gamma_{n+1} \supset \Gamma_n$ and all have P as one end point. Then γ is the outer limit of γ_n, where γ_n is the continuum corresponding to Γ_n. Now γ is not closed since Γ is not; hence if γ were entirely interior to d^*, it would have some limit points in $\bar{d}$ and from the continuity of $U(u)$, these would be carried into Q

which is impossible. Hence $\bar{\gamma}$ lies outside d^* so that all the points corresponding to P lie outside d^*. This proves (ii) since Q may be joined to ∞ by a continuum containing no points of R, and (iii) follows immediately from (ii).

Now it is well known that in any plane transformation all the points of the original set corresponding to an open set in the transformed set form a set open with respect to the original set. Thus, to any simply connected open region D, of R corresponds a connected open region d, of r. Now, let γ be a simple closed curve in d; $\Gamma = T(\gamma)$ lies in D. By (ii) and the simple connectivity of D, it follows that the interior of γ is carried into a subset of D. Thus d is simply connected.

The proof of (i) is the same if r and R are replaced by the Jordan regions $\bar{r}$ and $\bar{R}$ respectively. Let G be replaced by $\bar{G}$ and suppose a continuum $\bar{g}$, of $\bar{G}$ contains r^*. Now $\bar{r} - \bar{g}$ is obviously open, connected (by i), and is simply connected, since its boundary is a continuum clearly. Thus we may carry $\bar{r}$ in a continuous monotone way into the sphere, the transformation carrying $\bar{g}$ into a point being $1-1$ otherwise. This determines a collection $\bar{G}_1$ on the sphere which is topologically equivalent to $\bar{G}$, and of which no continuum separates the sphere again using (i). From a theorem of R. L. Moore (P. S. T., Theorem 19, chapter V), such a collection is topologically equivalent to the sphere. But $\bar{G}$ was given topologically equivalent to $\bar{R}$ which contradicts the above. Thus no continuum of $\bar{G}$ contains r^*. Having shown this, we may extend $\bar{T}$ to the whole of Π, $\bar{T}$ being $1-1$ for u outside $\bar{r}$. By letting r_1 be an open Jordan region, containing $\bar{r}$ and by considering the extended transformation in this region, we see that the remaining results may be easily shown.

THEOREM 2. *Let T, $T : U = U(u)$, $u \in r$, be a continuous monotone transformation of an open simply connected region r into another open simply connected region R. Then there exists a sequence $\{T_n\}$, $T_n : U = U_n(u)$, of $1-1$ continuous transformations of r into R such that the $U_n(u)$ converge uniformly to $U(u)$ on each closed subset of r. If we replace r and R by Jordan regions $\bar{r}$ and $\bar{R}$ respectively, we may choose the sequence $\{T_n\}$ so that the $U_n(u)$ converge uniformly on the whole of $\bar{r}$.*

Proof. It is clear that we may assume R to be the interior of Q, the square $(0,0;\ 1,1)$. Define $C_{i,j}$ as the line $V = j/2^i$, $D_{i,j}$ as the line $U = j/2^i$, $c_{i,j}$ as the connected set of r corresponding to $C_{i,j}$, and $d_{i,j}$ the connected set of r corresponding to $D_{i,j}$ $(i = 1, 2, \cdots;\ j = 1, 2, \cdots, 2^{i-1})$. Now, for each i, $\Gamma_i = r - \sum_{j=1}^{2^{i-1}} c_{i,j}$ and $\Delta_i = r - \sum_{j=1}^{2^{i-1}} d_{i,j}$ each consist of 2^i

mutually separated simply connected regions $\gamma_{i,j}$ and $\delta_{i,j}$ respectively $(j = 1, \cdots, 2^i)$; let $r^i_{j,k}$ be the simply connected region (corresponding to the open square $R^i_{j,k}$ of R) $\gamma_{i,j} \cdot \delta_{i,j}$. In each $r^i_{j,k}$ $(j, k = 1, \cdots, 2^i - 1)$ (i. e., omit the last row and column), choose a point $p^i_{j,k}$, and, for each j, $1 \leqq j \leqq 2^i - 1$, let $\tilde{c}_{i,j}$ be an open curve which passes through all the $p^i_{j,k}$ in $\gamma_{i,j}$, lies entirely interior to $\gamma_{i,j}$, divides it into exactly two parts, but is such that no segment of it divides $\gamma_{i,j}$; and for each k, $1 \leqq k \leqq 2^i - 1$, let $\tilde{d}_{i,k}$ be a similar curve in $\delta_{i,k}$ which has only the point $p^i_{j,k}$ in common with the curve $\tilde{c}_{i,j}$ for each j. Now it is clear that we may find a $1 - 1$ continuous transformation of r into R which carries $\tilde{c}_{i,j}$ into $C_{i,j}$ and $\tilde{d}_{i,j}$ into $D_{i,j}$. We define T_i, $T_i : U = U_i(u)$, to be any such transformation.

Now let p be any point of r, c_p the continuum of the collection corresponding to $U(u)$ containing it, and P the point of R corresponding to it under T. Let $C_1{}^i$, $C_2{}^i$, $D_1{}^i$, and $D_2{}^i$ be division lines of the i-th division of R $(= Q)$ such that, for each i, P is interior to the rectangle R_i cut off by the lines and such that $P = \prod_{i=1}^{\infty} R_i$. Let $c_j{}^i$, $d_j{}^i$, and r^i be the sets in r corresponding under T to $C_j{}^i$, $D_j{}^i$, and R_i, respectively $(j = 1, 2)$; and let $\tilde{c}^i_{n,j}$, $\tilde{d}^i_{n,j}$, and $\tilde{r}_{i,n}$ be the corresponding sets under T_n $(n \geqq i)$, the $\tilde{c}^i_{n,j}$ and $\tilde{d}^i_{n,j}$ being among the $\tilde{c}_{n,j}$ and $\tilde{d}_{n,j}$. Then, for each i, there exists an $N_{i,P}$ such that, for $n > N_{i,P}$, $p \in c_P \subset \tilde{r}_{i,n}$ and hence $P_n = T_n(p) \in R_i$. Hence the sequence $\{U_i(u)\}$ converges for each u. Now, if we restrict ourselves to the region $\bar{r}_n$ with its boundary which corresponds under T to the closed square $1/2^n \leqq u, v \leqq 1 - 1/2^n$, we see that the convergence is uniform. For it is clear that we can find a $\delta_i > 0$ such that if $| u_p - u_q | < \delta_i$, p and q $(p, q \in \bar{r}_n)$ are necessarily within some "cluster of four" of the $\bar{r}^i_{j,k}$ corresponding to four of the $\bar{R}^i_{j,k}$ in a cluster of the form $\boxplus$, and hence within a cluster of nine of the $\bar{r}^{i,m}_{j,k} = T_m{}^{-1}(\bar{R}^i_{j,k})$. Thus for $m > i$,

$$| U_{P_m} - U_{Q_m} | < 3(2)^{1/2}/2^i \quad \text{if} \quad | u_p - u_q | < \delta_i.$$

4. *The representation of surfaces on hemicactoids.* In this section, it is first shown that, if $\bar{H}$ is any hemicactoid and $X(U)$ is a continuous vector function defined on $\bar{H}$, (1) there exists a monotone transformation $U = U(u)$ of a Jordan region $\bar{r}$ into $\bar{H}$ which carries the subcollection G_0 of the corresponding collection G into $\bar{B}$, and (2) if we define $x(u) = X[U(u)]$, then $x = x(u)$ is a surface and all surfaces thus obtained are identical in the sense of Fréchet. This suggests the possibility of allowing surfaces to be represented on hemicactoids. The remainder of the section then treats of the fundamental questions of such a theory of surfaces.

We have defined the Fréchet distance $\| S_1, S_2 \|$ between two surfaces S_1 and S_2 in the introduction. The following lemma is an immediate consequence of these definitions.

LEMMA 1. (a) $\| S_1, S_2 \| = \| S_2, S_1 \|$; (b) $\| S_1, S_3 \| \leqq \| S_1, S_2 \| + \| S_2, S_3 \|$; (c) *if* $\| S_1, \bar{S}_1 \| = \| S_2, \bar{S}_2 \| = 0$, *then* $\| S_1, S_2 \| = \| \bar{S}_1, \bar{S}_2 \|$; *and* (d) *if* $\lim_{n\to\infty} \| S, S_n \| = 0$ *and* $x = x(u)$ *is any representation of* S, *we can find representations* $x = x_n(u)$, *of* S_n *such that the* $x_n(u)$ *converge uniformly to* $x(u)$.

Definition 1. If $\lim_{n\to\infty} \| S, S_n \| = 0$, we say that S_n *approaches* S or $\lim_{n\to\infty} S_n = S$.

LEMMA 2. *Let* T, $T : U = U(u)$, *and* T_1, $T_1 : U = U_1(u)$, *be two continuous monotone transformations of* $\bar{r}$ *into a base set* $\bar{B}$, *and* G *and* G_1 *the collections of maximal continua over each of which* $U(u)$ *and* $U_1(u)$, *respectively, is constant (hereafter, we shall call these collections, the collections corresponding to the given monotone transformations). If no continuum of either* G *or* G_1 *separates the plane and we define the "surfaces"* S *and* S_1 *by* $S : U = U(u)$, $S_1 : U = U_1(u)$, *then* S *and* S_1 *are identical in the sense of Fréchet.*

Proof. Let T and T_1 be extended monotonely to the whole of Π, each being $1 - 1$ outside of $\bar{r}$. Let $\bar{R}_n$ be a sequence of closed Jordan regions such that $\bar{R}_n \supset \bar{R}_{n+1}$ for each n and $\prod_{n=1}^{\infty} \bar{R}_n = \bar{B}$; and let $\bar{r}_n$ be $T^{-1}(\bar{R}_n)$ and $\bar{r}_{1n}$ be $T_1^{-1}(\bar{R}_n)$. It is clear that the Jordan closed curves r^*_n and r^*_{1n} approach r^* uniformly and hence also that $S_n \to S$ and $S_{1n} \to S_1$ respectively where S_n and S_{1n} are defined by

$$S_n : U = U(u),\ u \in \bar{r}_n, \qquad S_{1n} : U = U_1(u),\ u \in \bar{r}_{1n}.$$

But now, by Theorem 2, § 3, there exists a sequence $\{T_m\}$, $T_m : U = U_m(u)$, of $1 - 1$ continuous transformations of $\bar{r}_n$ into $\bar{R}_n$ which approaches T uniformly. Thus it follows immediately (from the definition of a surface) that $S_n = S'_n$, where $S'_n : U' = U, U \in \bar{R}_n$. Similarly $S_{1n} = S'_n$, so that $S_n = S_{1n}$, for each n. Thus it is easy to see that $S = S_1$.

LEMMA 3. *Let* S, $S : x = x(u)$, $u \in \bar{r}$, *be a surface and let* $\bar{g}$ *be a continuum of* $\bar{r}$ *which corresponds to a point of* S *and possesses a finite number of bounded (simply connected) complementary domains* $D_1, \cdots, D_n$. *Let* C_1 *be a simple closed curve in* D_1 *and* R_1 *the Jordan subregion of* D_1 *bounded by* C_1. *Let* T_1 *be a* $1 - 1$ *continuous transformation of* D_1 *into* R_1 *and let*

S_1 be represented by $x = x_1(\bar{u})$, $\bar{u} \in \bar{r}$, where $x_1(\bar{u}) = x(u)$, u and $\bar{u}$ being identical if u is outside $\bar{D}_1$, u and $\bar{u}$ corresponding under T_1 if $u \in D_1$, or $\bar{u} \in \bar{D}_1 - R_1$ and $u \in D^*_1$, simultaneously.

Then $x_1(\bar{u})$ is continuous and S_1 is a surface identical with S.

Proof. It is easily seen that $x_1(\bar{u})$ is continuous.

Let $R_1^{(m)}$ be a sequence of open Jordan regions where $R_1^{(m)} \subset R_1^{(m+1)}$ for each m and $D_1 = \sum_{m=1}^{\infty} R_1^{(m)}$. Let $T_1^{(m)}$ be a $1-1$ continuous transformation of $\bar{R}_1$ into $\bar{R}_1^{(m)}$ and suppose that the sequence $\{T_1^{(m)}\}$ converges uniformly to T_1^{-1} on every closed subregion of R_1. It is clear that we may define for each m, a $1-1$ continuous transformation of $\bar{D}_1 - R_1^{(m)}$ into $\bar{D}_1 - R_1$ which is the identity near D^*_1 and which preserves a preassigned correspondence between R^*_1 and $R^{*(m)}_1$. Hence we may define a sequence $\{T_m\}$, $T_m : \bar{u} = \bar{u}_m(u)$, of $1-1$ continuous transformations of $\bar{r}$ into itself which is the identity for u not in $\bar{D}_1$, is $(T_1^{(m)})^{-1}$ for $u \in \bar{R}_1^m$, and is the correspondence in the preceding sentence for $u \in \bar{D}_1 - R_1^{(m)}$. If we define $x_m(u) = x_1[\bar{u}_m(u)]$, it is easy to see that $\{x_m(u)\}$ converges uniformly to $x(u)$. But each of the surfaces S_m, $S_m : x = x_m(u)$, is identical with S_1 and by the preceding sentence, $\lim_{m\to\infty} \| S, S_m \| = 0$. Hence $S = S_1$.

LEMMA 4. *Let $\bar{C}$ be a double cone and $X(U)$ a continuous vector function defined on it. Suppose there exist two monotone transformations T, $T : U = U(u)$, and T_1, $T_1 : U = U_1(u)$, of $\bar{r}$ into $\bar{C}$, r^* being carried by each into the same endpoint of $\bar{C}$. Let S and S_1 be defined by*

$$S : x = x(u) = X[U(u)], \qquad S_1 : x = x_1(u) = X[U_1(u)].$$

Then $S = S_1$.

Proof. Let G and G_1 be the collections of continua in $\bar{r}$ corresponding to T and T_1 respectively and let g^* and g^*_1 be the continua of G and G_1, respectively, which contain r^*. Now, clearly (using Lemma 2, § 2 and the fact that g^* and g^*_1 are continua), the regions $D = \bar{r} - g^*$ and $D_1 = \bar{r} - g^*_1$ are simply connected. Now, let $U = U_2(v)$ be a monotone transformation of the unit circle $\bar{R}_0$ into $\bar{C}$ which carries R^*_0 into the above endpoint and is otherwise $1-1$ and let S_2 be represented on $\bar{R}_0$ by $x = \chi(v) = X[U_2(v)]$. Clearly there are "induced" monotone continuous transformations $v = v(u)$, and $v = v_1(u)$, of D and D_1, respectively, into R_0 $(= \bar{R}_0 - R^*_0)$ such that $U(u) = U_2[v(u)]$ and $U_1(u) = U_2[v_1(u)]$ for u in D and D_1 respectively. Now there exist sequences $\{\tau_n\}$ and $\{\tau_{1,n}\}$, $\tau_n : v = v_n(u)$, $\tau_{1,n} : v = v_{1,n}(u)$, of $1-1$ continuous transformations of D and D_1 respectively, into R_0, such

that $\{v_n(u)\}$ and $\{v_{1,n}(u)\}$ converge uniformly to $v(u)$ and $v_1(u)$ on all closed subsets of D and D_1. Now let $\{\bar{D}_m\}$ and $\{\bar{D}_{1m}\}$ be sequences of closed Jordan regions such that $\bar{D}_m \subset \bar{D}_{m+1}$ and $\bar{D}_{1m} \subset \bar{D}_{1,m+1}$ for each m, and $D = \sum_{m=1}^{\infty} \bar{D}_m$, $D_1 = \sum_{m=1}^{\infty} \bar{D}_{1m}$. Now, let S_m and S_{1m} be defined by

$$S_m : x = x(u),\ u \in \bar{D}_m, \qquad S_{1m} : x = x_1(u),\ u \in \bar{D}_{1m}.$$

Clearly $\{S_m\}$ and $\{S_{1m}\}$ converge to S and S_1, respectively. Now, for each m, let n_m be so large that $|\, v_{n_m}(u) - v(u)\,| < 1/m$, $u \in \bar{D}_m$; and

$$|\, v_{1,n_m}(u) - v_1(u)\,| < 1/m,\ u \in \bar{D}_{1,m},$$

and let $\bar{R}_m$ and $\bar{R}_{1m}$ be the Jordan regions in R_0 corresponding, under τ_{n_m} to $\bar{D}_m$ and $\bar{D}_{1m}$, respectively. Let S'_m and S'_{1m} be defined by

$$S'_m : x = x_m(u) = X\{U_2[v_{n_m}(u)]\},\ u \in \bar{D}_m,$$
$$S'_{1m} : x = x_{1m}(u) = X\{U_2[v_{1,n_m}(u)]\},\ u \in \bar{D}_{1m}.$$

Since $x(u) = X\{U_2[v(u)]\}$ and $x_1(u) = X\{U_2[v_1(u)]\}$, it is easy to see that S_m' and S_m approach the same limit. Since $S_m' = S_m''$, $S_m'' : x = X[U_2(v)]$, $v \in \bar{R}_m$, and since it is clear that R^*_m converges to R^*_0, we see that $\{S'_m\}$ converges to S_2. Similarly $\{S_{1m}\}$ and $\{S'_{1m}\}$ converge to the same limit, $\{S'_{1m}\}$ converging to S_2. Thus $S = S_2 = S_1$, which proves the lemma.

LEMMA 5. *Let the hypotheses of Lemma* 4 *be satisfied with* $\bar{C}$ *being any simple cyclic chain of type A instead of merely a double cone. Then also,* $S_1 = S$.

Proof. First of all, let P and Q be two cut points of any simple cyclic chain $\bar{C}'$, of type A and suppose P separates Q from an endpoint E, of $\bar{C}'$. Let T' be a continuous monotone transformation of $\bar{r}$ into $\bar{C}'$ in which r^* is carried into E, and let $\bar{G}'$ be the corresponding collection of continua in $\bar{r}$. Now, using Lemma 2, § 2, we see that (1) the continua g'_P and g'_Q of G' corresponding to P and Q respectively, each bound exactly one simply connected region, say d'_P and d'_Q, (2) $g'_Q \subset d'_P$, (3) the region $r'_{P,Q} = d'_P - g'_Q - d'_Q$, between g'_P and g'_Q, is doubly connected, and (4) $g'_P + r'_{P,Q} + g'_Q$ can be mapped in a monotone continuous way on to the surface of a sphere, g'_P and g'_Q being carried into any two distinct points, the mapping being $1-1$ and continuous otherwise. Hence, if we replace the part of $\bar{C}'$ between P and Q by another simple cyclic chain $\bar{C}''_{P,Q}$ of type A with endpoints P and Q, we can find an upper semicontinuous collection, $G''_{P,Q}$ filling up $r'_{P,Q}$ so that $\bar{G}''_{P,Q} = G''_{P,Q} + g'_P + g'_Q$ and $\bar{G}' - G'_{P,Q} + G''_{P,Q}$ are upper semicontinuous collections topologically equivalent to $\bar{C}''_{P,Q}$ and $\bar{C}' - \bar{C}'_{P,Q} + \bar{C}''_{P,Q}$ respec-

tively, where $C'_{P,Q}$ is the part of $\bar{C}'$ between P and Q, $\bar{C}'_{P,Q} = C'_{P,Q} + P + Q$, and $G'_{P,Q}$ is the subset of G' filling up $\gamma'_{P,Q}$ and corresponding to $C'_{P,Q}$. This can be done in virtue of (4) above and the fact that a collection $G''_{P,Q}$ can be found on the sphere topologically equivalent to $\bar{C}''_{P,Q}$ in which P and Q correspond to points on the sphere (using Lemma 9, § 2).

Hence, form $\bar{C}_m$ from $\bar{C}$ by replacing each double cone of $\bar{C}$ of length $< 1/m$ by its axis, and define $X_m(U)$ on $\bar{C}_m$ by letting $X_m(U) = X(U)$, $U \in \bar{C} \cdot \bar{C}_n$, and $X_m(U) = X(U')$, for U on the axis of one of the replaced double cones of $\bar{C}$, U' ranging linearly over the broken line consisting of the two equal sides of a generating triangle of that double cone as U ranges over the axis. In each region in $\bar{r}$ corresponding to such a double cone, replace the part of $\bar{G}$ and $\bar{G}_1$ corresponding to the double cone by a new collection topologically equivalent to the axis. It is easily seen that the new $\bar{G}_m$ and $\bar{G}_{1m}$ are upper semicontinuous. Hence let $U = U_m(u)$ and $U = U_{1m}(u)$ be the corresponding monotone transformations of $\bar{r}$ into $\bar{C}_m$; and let S_m and S_{1m} be defined by

$$S_m : x = x_m(u) = X_m[U_m(u)], \quad S_{1m} : x = x_{1m}(u) = X_m[U_{1m}(u)].$$

Then it is clear that $\lim_{m\to\infty} S_m = S$ and $\lim_{m\to\infty} S_{1m} = S_1$, noticing that $\{U_m(u)\}$ and $\{U_{1m}(u)\}$ converge uniformly to $U(u)$ and $U_1(u)$, respectively. Now, let $\bar{C}_{m,n}$ be formed from $\bar{C}_m$ by dividing each interval of $\bar{C}_m$ between successive double cones into n equal parts and replacing each subinterval by a double cone with the same end points and the proper vertex angle. Define $X_{m,n}(U)$ on $\bar{C}_{m,n}$ by the relations

$$X_{m,n}(U) = X_m(U), U \in \bar{C}_m \cdot \bar{C}_{m,n}, \; X_{m,n}(U) = X_m(U'), U \in \bar{C}_{m,n} - \bar{C}_m \cdot \bar{C}_{m,n}$$

U' being the intersection, with the axis of $\bar{C}_{m,n}$ of a plane through U and perpendicular to the axis. Now let $U = U_{m,n}(u)$ and $U = U_{1,m,n}(u)$ be monotone transformations of $\bar{r}$ into $\bar{C}_{m,n}$ coinciding with $U_m(u)$ and $U_{1m}(u)$ except on regions corresponding to $\bar{C}_m - \bar{C}_m \cdot \bar{C}_{m,n}$ and define $S_{m,n}$ and $S_{1,m,n}$ as indicated above. Now, again, $\lim_{n\to\infty} S_{m,n} = S_m$ and $\lim_{n\to\infty} S_{1,m,n} = S_{1,m}$. Hence we shall prove our theorem for chains $\bar{C}$ made up entirely of a finite number of abutting double cones from which the general case will follow by limiting processes.

We shall prove the theorem for this case by induction on the number n, of double cones. We have proved the theorem when $n = 1$ in Lemma 4. Suppose the theorem is true for $n = k$ and that $\bar{C}$ has $k + 1$ double cones. Let $\bar{C}_k$ consist of the first k double cones of $\bar{C}$, starting from E, let E_k be the other end point of $\bar{C}_k$, and let g_k and g_{1k} be the continua of $\bar{r}$ corresponding

to E_k under T and T_1, respectively and d_k and d_{1k} the regions bounded, respectively, by g_k and g_{1k}. Let C_k be a simple closed curve in d_k and C_{1k} one in d_{1k} and let R_k be the Jordan subregion of d_k bounded by C_k and R_{1k} that of d_{1k} bounded by C_{1k}. Let T_k be a $1-1$ continuous transformation of d_k into R_k and T_{1k} one of d_{1k} into R_{1k}. Let $\bar{x}(\bar{u})$ and $\bar{x}_1(\bar{u}_1)$ be defined for $\bar{u}$ and $\bar{u}_1$ on $\bar{r}$ by the relations $\bar{x}(\bar{u}) = x(u)$ and $\bar{x}_1(\bar{u}_1) = x_1(u)$, where $u = \bar{u}$, $u \in \bar{r} - \bar{d}_k$, u and $\bar{u}$ correspond under T_k if $u \in d_k$, or $u \in d^*_k$ and $\bar{u} \in \bar{d}_k - R_k$, simultaneously, an analogous statement holding for u and $\bar{u}_1$. Then $x = \bar{x}(\bar{u})$ and $x = \bar{x}_1(\bar{u}_1)$, $\bar{u}$, $\bar{u}_1 \in \bar{r}$, are representations of S and S_1, respectively, by Lemma 3. Now, let S_k and S_{1k} be given respectively by $x = x_k(u)$ and $x = x_{1k}(u)$, where

$$x_k(u) = x(u),\ u \in \bar{r} - d_k, \qquad x_k(u) = X(U_{E_k}),\ u \in d_k,$$

and $x_{1k}(u)$ is defined analogously. By Lemma 3, S_k and S_{1k} are also given by $x = \bar{x}_k(\bar{u})$ and $x = \bar{x}_{1k}(\bar{u}_1)$, respectively, where $\bar{x}_k(\bar{u})$ and $\bar{x}_{1k}(\bar{u}_1)$ are obtained from $\bar{x}(\bar{u})$ and $\bar{x}_1(\bar{u}_1)$ in the way that $x_k(u)$ and $x_{1k}(u)$ are obtained from $x(u)$ and $x_1(u)$, $\bar{r} - d_k$ and $\bar{r} - d_{1k}$ being replaced by $\bar{r} - R_k$ and $\bar{r} - R_{1k}$, respectively. By hypothesis of the induction $S_k = S_{1k}$, so that there exists a sequence $\{\bar{\tau}_{k,n}\}$ of $1-1$ continuous transformations of $\bar{r}$ into itself, and hence a sequence $\{\tau_{k,n}\}$, $\tau_{k,n} : \bar{u}_1 = \bar{\bar{u}}_1^{(n)}(\bar{u})$, of such transformations of $\bar{r} - R_k$ into $\bar{r} - R_{1k}$ conserving a given correspondence between C_k and C_{1k} and such that the sequence $\{\bar{x}_n(\bar{u})\}$, defined on $\bar{r} - R_k$ by $\bar{x}_n(\bar{u}) = \bar{x}_{1k}[\bar{\bar{u}}_{1n}(\bar{u})]$, converges uniformly on $\bar{r} - R_k$ to $\bar{x}(\bar{u})$. Now let σ_k and σ_{1k} be the surfaces defined on $\bar{r}$ by $x = x'_k(u)$ and $x = x'_{1k}(u)$, respectively, where

$$x'_k(u) = x(u),\ u \in d_k, \qquad x'_k(u) = X(U_{E_k}),\ u \in \bar{r} - d_k,$$

and $x'_{1k}(u)$ is defined analogously. By Lemma 3, σ_k and σ_{1k} are represented on $\bar{R}_k$ and $\bar{R}_{1k}$, respectively, by $x = \bar{x}(\bar{u})$, $\bar{u} \in \bar{R}_k$, and $x = \bar{x}_1(\bar{u}_1)$, $\bar{u}_1 \in \bar{R}_{1k}$. By Lemma 4, $\sigma_k = \sigma_{1k}$ (the part of S and S_1 corresponding to the last double cone of $\bar{C}$) so there exists a sequence $\{\tau'_{k,n}\}$, $\bar{u}_1 = \bar{u}'^{(n)}_1(\bar{u})$, of $1-1$ continuous transformations of $\bar{R}_k$ into $\bar{R}_{1k}$ preserving the above correspondence between R^*_k and R^*_{1k}, such that the sequence $\{\bar{x}_n(\bar{u})\}$, defined on $\bar{R}_k$ by $\bar{x}_n(\bar{u}) = \bar{x}_1[\bar{u}'^{(n)}(\bar{u})]$, converges uniformly on $\bar{R}_k$ to $\bar{x}(\bar{u})$. Now, for each n, $\tau_{k,n}$ and $\tau'_{k,n}$ unite to form a $1-1$ continuous transformation $T_{k,n}$, $T_{k,n} : \bar{u}_1 = \bar{u}_1^{(n)}(\bar{u})$, of $\bar{r}$ into itself and the above $\bar{x}_n(\bar{u})$ is seen to be continuous over $\bar{r}$ and also given by $\bar{x}_n(\bar{u}) = \bar{x}_1[\bar{u}_1^{(n)}(\bar{u})]$. Hence $S = S_1$ and our induction is complete.

THEOREM 1. *Let $\bar{H}$ be a hemicactoid and $X(U)$ a continuous vector function defined on it. There exists a continuous monotone transformation $U = U(u)$, with associated upper semicontinuous collection G, of a Jordan*

region $\bar{r}$, into $\bar{H}$ which carries the subcollection G_0, of Lemma 8, § 2, topologically into $\bar{B}$. If we define $x(u) = X[U(u)]$, then $x = x(u)$ is a surface and all surfaces obtained in this way are identical.

Proof. The first part of the theorem has been proved in Theorem 4, § 2.

To demonstrate the last part, let T, $T: U = U(u)$, and T_1, $T_1: U = U_1(u)$, $u \in \bar{r}$, be two continuous monotone transformations with corresponding collections G and G_1, which carry the above subcollections G_0 and G_{10}, into $\bar{B}$. Let S and S_1 be the corresponding surfaces obtained as above. Let us represent $\bar{H}$ by

$$\bar{H} = \bar{B} + \sum_{n=1}^{\infty} \bar{C}_n + H^*,$$

where, for each n, $\bar{C}_n$ is a simple cyclic chain of type A such that $\bar{C}_n \cdot [\bar{B} + \sum_{j=1}^{n-1} \bar{C}_j]$ is an end point of $\bar{C}_n$, and H^* is a completely disconnected set of limit points of $\bar{B} + \sum_{j=1}^{\infty} \bar{C}_n$. Let $\bar{H}_0 = \bar{B}$, $\bar{H}_n = \bar{B} + \sum_{j=1}^{n} \bar{C}_n$, let G_n and G_{1n} be the subcollections of G and G_1, respectively, corresponding (under T and T_1) to $\bar{H}_n$, and let $\hat{G}_n$ and $\hat{G}_{1n}$ denote the respective sets of points of $\bar{r}$ covered by these collections. Now it is easy to see (using the method of Lemma 8, § 2) that $\hat{G}_n$ is a continuum, that every point of $\bar{r}$ either belongs to $\hat{G}_n$ or is in a bounded complementary domain of a continuum of $\hat{G}_n$, and that the same statements hold for $\hat{G}_{1n}$. Thus, for each $n \geqq 0$, let G'_n be the collection of continua g'_n, each g'_n being obtained from a g_n by adding to it all the complementary domains of $\hat{G}_n$ which it bounds, and let G'_{1n} be similarly formed from G_{1n}. Let T_n, $T_n: U = U_n(u)$, and T_{1n}, $T_{1n}: U = U_{1n}(u)$, be the monotone continuous transformations of $\bar{r}$ into $\bar{H}_n$, which have G'_n and G'_{1n} as their respective corresponding collections. Let S_n and S_{1n} be defined by

$$S_n : x = x_n(u) = X[U_n(u)], \qquad S_{1n} : x = x_{1n}(u) = X[U_{1n}(u)].$$

Then clearly the sequences $\{U_n(u)\}$ and $\{U_{1n}(u)\}$, converge uniformly to $U(u)$ and $U_1(u)$ respectively, and $\lim_{n\to\infty} S_n = S$ and $\lim_{n\to\infty} S_{1n} = S_1$. Hence, if we prove the theorem for the case when the number of $\bar{C}_n$ is finite, the general case will follow by a simple limit process. But the proof of this finite case may be handled by induction in precisely the same way that the corresponding proof in Lemma 5 was carried through.

The above theorem gives a logical justification for surfaces to be "represented" on hemicactoids. The remainder of this section will be devoted to certain essential parts of a theory of such representations.

Definition 2. The two surfaces $x = X(U)$, $U \in \bar{H}$, and $x = x(u)$, $u \in \bar{r}$,

$\bar{H}$ being a hemicactoid, will be said to be *identical* if there exists a continuous monotone transformation $U = U(v)$, $v \in \bar{R}$, of a Jordan region $\bar{R}$ into $\bar{H}$, with corresponding collection G, carrying G_0 topologically into $\bar{B}$, and such that the surface $x = \xi(v) = X[U(v)]$, $v \in \bar{R}$, is identical with the surface $x = x(u)$, $u \in \bar{r}$. By the above theorem, all the surfaces $x = \xi(v)$ thus obtained are identical.

Definition 3. The vector function $X(U)$, $U \in \bar{H}$, $\bar{H}$ a continuum, is *non-degenerate* on $\bar{H}$ if $X(U)$ is not constant over any continuum of $\bar{H}$ containing two points.

Definition 4. The representation $x = X(U)$, $U \in \bar{H}$, of a surface on the hemicactoid $\bar{H}$, is *non-degenerate* if $X(U)$ is non-degenerate on $\bar{H}$.

THEOREM 2. *Any surface S, $S : x = x(u)$, $u \in \bar{r}$, can be represented non-degenerately on some hemicactoid. If $x = X(U)$, $U \in \bar{H}$, and $x = X_1(U_1)$, $U_1 \in \bar{H}_1$, are two such representations of S, then there exists a $1-1$ continuous transformation $U_1 = U_1(U)$, of $\bar{H}$ into $\bar{H}_1$ carrying $\bar{B}$ into $\bar{B}_1$ such that $X_1[U_1(U)] = X(U)$.*

Proof. Let G be the upper semicontinuous collection of maximal continua over each of which $x(u)$ is constant. Now, by Theorem 3, § 2, we may carry G topologically into a hemicactoid $\bar{H}$, by a monotone continuous transformation $U = U(u)$, which also carries G_0 into $\bar{B}$, and if $U_1 = U_1(u)$ is another such transformation, carrying G topologically into a hemicactoid $\bar{H}_1$ and G_0 into $\bar{B}_1$ then there exists a $1-1$ continuous transformation $U_1 = \bar{U}_1(U)$, of $\bar{H}$ into $\bar{H}_1$ and $\bar{B}$ into $\bar{B}_1$ such that $\bar{U}_1[U(u)] = U_1(u)$. Hence if we define $X(U)$ on $\bar{H}$ so that $X[U(u)] = x(u)$ as can obviously be done, we have only to show that $X(U)$ is continuous and non-degenerate in order to prove the theorem. But this is immediate, for, let $\{U_n\}$, $U_n \in \bar{H}$, converge to U, and let $\{g_n\}$ and g be the corresponding continua in $\bar{r}$. Then by Theorem 2, § 2, g is the limit continuum of $\{g_n\}$ and there exists a sequence of points $\{p_n\}$, $p_n \in g_n$, such that every convergent subsequence converges to a point p of g; this proves the continuity of $X(U)$. The non-degeneracy is obvious.

Definition 5. A point P, of a hemicactoid $\bar{H}$, is said to be separated from the point ∞ in the plane Π, of $\bar{B}$ by a set σ if every connected set in $\bar{H}$ containing P and a point of $(\Pi - \bar{B})^*$ also contains a point of σ.

THEOREM 3. *If $x = X(U)$, $U \in \bar{H}$, is a non-degenerate representation of S, and $x = x(u)$, $u \in \bar{h}$, is any other representation of S ($\bar{H}$ and $\bar{h}$ hemicactoids) there exists a continuous monotone transformation $U = U(u)$, with corresponding collection of continua $G^{(\bar{h})}$, which carries $\bar{h}$ into $\bar{H}$, is such that*

$X[U(u)] = x(u)$, and carries the collection $G_0^{(\bar{h})}$, topologically into $\bar{B}$, $G_0^{(\bar{h})}$ consisting of the continua of $G^{(\bar{h})}$, no point of any of which is separated from the point ∞ in the plane of $\bar{b}$ by a continuum of $G^{(\bar{h})}$. If $x = x(u)$ is also non-degenerate, any such transformation is $1-1$. The collection $G^{(\bar{h})}$ may be taken to be the collection of maximal continua of $\bar{h}$ over each of which $x(u)$ is constant.

Proof. The last two statements are consequences of Theorem 2.

Now take $G^{(\bar{h})}$ as the collection of maximal continua of $\bar{h}$ over each of which $x(u)$ is constant and let $u = u(v)$ be a continuous monotone transformation with corresponding collection Γ, which carries $\bar{r}$ into $\bar{h}$ and Γ_0 topologically into $\bar{b}$. Now it is easy to verify the fact that $G^{(\bar{h})}$ is carried topologically into the collection G, of maximal continua over each of which $\xi(v) = x[u(v)]$ is constant. Now it is clear that $u = u(v)$ sets up a topological correspondence between the points of $(\Pi - \bar{B})^*$ and those continua of Γ which have points in common with r^*, that each continuum of G corresponds to a continuum of elements of Γ, that $u = u(v)$ carries a connected point set of $\bar{r}$ into one on $\bar{h}$, and that, to a connected set of points of $\bar{h}$ corresponds one such in $\bar{r}$. Hence it follows easily that $u = u(v)$ establishes a topological relation between the continua of $G_0^{(\bar{h})}$ and G_0. Now, let $U_1 = U_1(v)$ be a continuous monotone transformation with corresponding collection G, which carries $\bar{r}$ into $\bar{H}_1$ and G_0 topologically into $\bar{B}_1$, and let $X_1(U_1)$ be determined (see Theorem 2) so $X_1[U_1(v)] = \xi(v)$. Then, by Definition 2, S is given by $x = \xi(v)$, $v \varepsilon \bar{r}$, $x = X_1(U_1)$, $U_1 \varepsilon \bar{H}_1$, and by hypothesis also by $x = X(U)$, $U \varepsilon \bar{H}$. Hence there exists a $1-1$ continuous transformation $U = U(U_1)$, of $\bar{H}_1$ into $\bar{H}$ and $\bar{B}_1$ into $\bar{B}$ such that $X[U(U_1)] = X_1(U_1)$. By these successive transformations we have set up a topological relation $U = U(u)$, of the continua of $G^{(\bar{h})}$ into the points of $\bar{H}$ which is easily seen to possess the required properties.

Definition 6. The transformations $U = U(u)$, obtained this way will be said to be *induced* by the change of representation.

5. *On non-degenerate surfaces.* In this section, we consider a very important class of surfaces, which from some points of view are the most general surfaces to which one would like to apply the name of "surface." A characterization of such surfaces in terms of their given representation is developed and an interesting subclass of such surfaces is pointed out.

Definition 1. A surface is said to be *non-degenerate* if it possesses a non-degenerate representation on a Jordan region.

THEOREM 1. *If $x(u)$ is non-degenerate on a Jordan region $\bar{r}$, and $u = u(U)$ is a $1-1$ continuous transformation of a Jordan region $\bar{R}$ into $\bar{r}$ and we define $X(U) = x[u(U)]$, then $X(U)$ is non-degenerate on $\bar{R}$.*

Remark. In the case of a non-degenerate surface, it is clear that any hemicactoid on which the surface may be represented non-degenerately consists only of its base set and this of a single rhombus.

THEOREM 2. *If* (i) S, $S : x = X(U)$, *and* S_n, $S_n : x = X_n(U)$, $U \in \bar{R}$, *are all non-degenerate surfaces,* (ii) $\lim_{n\to\infty} \| S, S_n \| = 0$, *and* (iii) $x = x(u)$, $u \in \bar{r}$, *is a non-degenerate representation of S, we can find a sequence of non-degenerate representations $x = x_n(u)$, of S_n so that the $x_n(u)$ approach $x(u)$ uniformly.*

Proof. Each S_n possesses a non-degenerate representation $x = \tilde{x}_n(\tilde{U})$, $\tilde{U} \in \bar{\bar{R}}$, on some Jordan region $\bar{\bar{R}}$. From the definition of Fréchet distance, it follows that, for each n, we can find a $1-1$ continuous transformation $\tilde{U} = \tilde{U}_n(u)$, of $\bar{r}$ into $\bar{\bar{R}}$ such that

$$| x(u) - x_n(u) | < 2 \| S, S_n \|, \qquad x_n(u) = \tilde{x}_n[\tilde{U}_n(u)].$$

From this and Theorem 1, it is clear that $x = x_n(u)$ is the desired sequence.

THEOREM 3. *A necessary and sufficient condition that the surface S, $S : x = x(u)$, $u \in \bar{r}$, be non-degenerate is that* (i) *no continuum over which $x(u)$ is constant separates $\bar{r}$ and* (ii) *$x(u)$ is not constant over r^*.*

Proof. If S is non-degenerate, there exists a monotone continuous transformation $U = U(u)$, *of* $\bar{r}$ into the square $\bar{Q}$ which carries the collection G, of maximal continua of $\bar{r}$ over each of which $x(u)$ is constant topologically into the points of $\bar{Q}$ (using Theorem 3, § 4). Now by Theorem 1, § 3, no continuum of G separates $\bar{r}$ or contains all of r^*.

Now suppose these conditions are satisfied by the continua of G. Then we can make a monotone continuous transformation $u_1 = u_1(u)$, of $\bar{r}$ into $\bar{r}_1$ which is $1-1$ inside r and carries the intervals (may reduce to points) of r^* over each of which $x(u)$ is constant into points of r^*_1. The collection G, is carried topologically into a collection G_1, no continuum of which has more than one point in common with r^*_1, since $x_1(u_1) = x(u)$ is not constant over any interval of r^*_1 and if a g_1 contained two points of r^*_1 it would separate $\bar{r}_1$ and the corresponding g would separate $\bar{r}$. It is clear that $x = x_1(u_1)$, $u_1 \in \bar{r}_1$, is a representation of S. Let G'_1 be the upper semicontinuous collection consisting of G_1 plus all the points of $\Pi - \bar{r}_1$. There exists † a continuous mono-

† R. L. Moore, *loc. cit.* (not P. S. T.).

tone transformation $U = U_1(u_1)$, which carries the continua of G'_1 topologically into the points of II. It is clear that r^*_1 is carried into a Jordan curve, which we may take to be Q^*, by this transformation. By Theorem 1, § 3, every point inside r_1 is carried inside or on Q^* and every point of r^*_1 we know is carried into Q^*. Thus $U = U_1(u_1)$ carries $\bar{r}_1$ into $\bar{Q}$ for if a point of Q were missed it would be necessary for points of r^*_1 to be carried into the interior of Q. Thus $U = U(u) = U_1[u_1(u)]$ carries $\bar{r}$ in a monotone way into $\bar{Q}$, carrying the continua of G topologically into the points of $\bar{Q}$. By arguments already used, we see that we can define $X(U)$ on $\bar{Q}$ so as to be continuous and non-degenerate, and so that $X[U(u)] = x(u)$. Since $U = U(u)$ is the uniform limit of a sequence of $1-1$ continuous transformations of $\bar{r}$ into $\bar{Q}$, it is easily seen that $x = X(u)$, $U \in \bar{Q}$ is a non-degenerate representation of S.

Definition 2. We say that the surface S, $S : x = x(u)$, $u \in \bar{r}$, is *bounded by a Jordan curve* C, if $x = x(p)$, $p \in r^*$, is a representation of C on r^*, i. e., if it is possible to find a continuous monotone transformation $P = T(p)$, of r^* into another simple closed curve R^*, such that there exists a continuous function $X(P)$, defined on R^* such that (i) $X[T(p)] = x(p)$ and (ii) the equation $x = X(P)$ gives a $1-1$ continuous representation of C on R^*.

Remark. It is easy to see that in this case, the base set of a hemicactoid on which S is represented non-degenerately, consists of a single rhombus.

Definition 3. A function $f(x, y)$, defined on a region $\bar{r}$, is said to be *monotone in the sense of Lebesgue* if, for every subregion d, of $\bar{r}$, the oscillation of $f(x, y)$ on $\bar{d}$ is equal to its oscillation on d^*.

Definition 4. The surface S, $S : x = x(u)$, $u \in \bar{r}$, is said to be *Lebesgue monotone* if each of the component functions, $x^i(u)$, is monotone in the sense of Lebesgue. It is easily seen that this property of a surface is invariant under changes of representation on Jordan regions.

THEOREM 4. *Any hemicactoid on which a Lebesgue monotone surface may be represented non-degenerately consists only of its base set.*

Proof. This is immediate, for no continuum over which $x(u)$ is constant can separate the plane so that the collection G of maximal continua of $\bar{r}$ over each of which $x(u)$ is constant is equal to its subcollection G_0.

THEOREM 5. *A monotone surface bounded by a Jordan curve C, is non-degenerate.*

Proof. This follows immediately from Theorem 4 and the remark after Definition 2.

THE RICE INSTITUTE.

CONCERNING SPHERICAL SPACES.†

By LEONARD M. BLUMENTHAL.‡

1. *Introduction.* The n-dimensional spherical space $S_{n,r}$ consists of the points of the surface of an $(n+1)$-dimensional sphere of radius r in a euclidean space of $n+1$ dimensions, with the distance between two points defined as geodesic (great circle) distance. A recent paper characterized those semi-metric spaces that are congruent (i. e., isometric) with a subset of the $S_{n,r}$.§ If $p_1, p_2, \cdots, p_n$ are n points of a semi-metric space and we form the symmetric determinant $\Delta(p_1, p_2, \cdots, p_n) = |\cos p_i p_j / r|$, $(i, j = 1, 2, \cdots, n)$, Theorems I_n, II_n, III_n of the paper referred to may be combined to give the following theorem.

THEOREM.¶ *A semi-metric space is congruent with a subset of $S_{n,r}$ if and only if, no pair of points of the space has a distance exceeding πr and,*

(i) *for every integer $k \leq n+1$ and every k points of the space, the determinant $\Delta(p_1, p_2, \cdots, p_k)$ is positive or zero,*

(ii) *the determinant Δ formed for every $n+2$ points of the space vanishes,*

(iii) *the determinant of every $n+3$ points vanishes.*

This theorem gives, essentially, the conditions that $n+3$ points of a semi-metric space be isometric with $n+3$ points of the $S_{n,r}$, and no assumptions of a qualitative nature are made concerning the semi-metric space containing the $n+3$ points.

The object of this paper is to show that if certain simple qualitative conditions are placed upon the semi-metric space, the chain of determinant conditions contained in hypothesis (i) of the theorem are implied by conditions involving only sets of *four points* of the space, while hypothesis (iii) is superfluous, being a consequence of hypothesis (ii). A semi-metric space satisfying our postulates is shown to be congruent with $S_{n,r}$ or with the spherical Hilbert space of radius r according as the dimension of the space is finite or infinite.

It is to be remarked that the method used in proving the fundamental

† Presented to the American Mathematical Society, March 30, 1934.

‡ National Research Fellow.

§ Blumenthal and Garrett, "Characterization of spherical and pseudo-spherical sets of points," *American Journal of Mathematics*, vol. 55 (1933), pp. 619-640.

¶ This result has also been obtained by L. Klanfer, of Vienna, a brief account of whose work appears in the *Ergebnisse Eines Mathematischen Kolloquiums*, Wien, Heft 4 (1933), p. 43.

Theorem 3.3 makes no use of the determinant conditions that express the necessary and sufficient conditions that $n+1$ points of a semi-metric space be congruent with $n+1$ points of $S_{n,r}$. Hence the procedure is applicable to those spaces for which analogous determinant conditions are unknown. *The corresponding theorem is valid for a wide class of n-dimensional spaces admitting a rotation about an* $(n-2)$*-dimensional subset.*†

2. A set of undefined elements, for suggestiveness called points, and a positive number r form an S_r space provided that to each pair p_1, p_2 of elements there corresponds a real number p_1p_2, called their distance, satisfying the following postulates:

Postulate A. $p_1p_2 = 0$, if and only if $p_1 = p_2$.

Postulate B. Every three elements p_1, p_2, p_3 of the space satisfy the triangle inequality; i. e., $p_1p_2 + p_1p_3 \geq p_2p_3$.

Postulate C. For every three elements p_1, p_2, p_3 of the space, the sum of the three distances they determine does not exceed $2\pi r$.

Postulate D. If p_1, p_2, p_3, p_4 are four elements of the space such that some triple contained in the four points is isometric with three points of a circle $S_{1,r}$, of radius r, then the four points are isometric with four points of a two-dimensional spherical surface $S_{2,r}$, of radius r.‡

From Postulates A, B it follows that $p_1p_2 \geq 0$, and $p_1p_2 = p_2p_1$, while it is evident from Postulates B, C that the distance of each two points of the space does not exceed πr. If the distance of two points of the space equals πr, the points are said to be diametral.

A set of points which is congruent § or isometric with a subset of $S_{n,r}$ is called r-spheric (S_n). It is clear that every three points of S_r is isometric with three points of $S_{2,r}$.

LEMMA 2.1. *Let p, q be any two non-diametral points of S_r. There is at most one point x of S_r such that $px + xq = pq$ and $px = c \cdot pq$, c, a constant, $0 < c < 1$.*

---

† I wish to express my thanks to Dr. Kurt Gödel for numerous suggestions given me during the preparation of this paper.

‡ This postulate may be stated arithmetically in terms of the vanishing of the determinant Δ of the four points provided any one of its third-order principal minors vanish.

The space S_r may also be defined as a *semi-metric* space such that (i) no distance exceeds πr, (ii) the determinant Δ formed for each triple of points is non-negative, and (iii) each set of four points of the space satisfies Postulate D. This definition of S_r does not involve explicitly the triangle inequality.

§ Two semi-metric sets M, M' are congruent or isometric if they may be mapped isometrically upon each other. We write $M \approx M'$.

Proof. Suppose x, x^* are two points of S_r which satisfy the conditions of the lemma, and consider the four points p, x, x^*, q. The triple p, x, q is clearly isometric with three points of $S_{1,r}$ and hence, by Postulate D there are four points p', x', $x^{*\prime}$, q' of $S_{2,r}$ isometric with the points p, x, x^*, q. But, by hypothesis x and x^* are each between † p and q; hence x' and $x^{*\prime}$ are each between p' and q' and the points p', x', $x^{*\prime}$, q' lie on $S_{1,r}$, since $p'q' = pq < \pi r$. Evidently $x' = x^{*\prime}$; then $x = x^*$ and the lemma is proved.

A semi-metric space is called *convex* provided that for each two distinct points of the space, the space contains at least one point between them. Now Menger has shown that if p, q are any two distinct points of a convex, complete, metric space R, then R contains *at least* one point x that satisfies the conditions of Lemma 2. 1. Since S_r is a metric space we have, combining Lemma 2. 1 and the theorem cited above, the following theorem:

THEOREM 2. 1. *If p, q are any two distinct non-diametral points of S_r and if S_r is convex and complete, then there is exactly one point x of S_r such that $px + xq = pq$ and $px = c \cdot pq$, c, a constant, $0 < c < 1$.*

Letting c take on all values between zero and one, we call the set of points $p + q + \{x\}$, the arc pq. Then evidently *two distinct non-diametral points of the convex, complete space S_r determine exactly one arc joining them,* and this arc is congruent with an arc of length pq of the circle $S_{1,r}$. The arc consists of the two points together with all the points of S_r between them. If two points of S_r are diametral, every point of S_r distinct from them lies between them and hence, in general, two such points do not determine an arc.

We suppose, further, that to each point p of S_r there corresponds a point $\bar{p}$ of S_r diametral to p; that is, $p\bar{p} = \pi r$. A space S_r having this property is said to be *diameterized.* It is clear from Postulate C that S_r can contain at most one point $\bar{p}$ diametral to p. Consider now two non-diametral points p, q of S_r and let $\bar{p}$, $\bar{q}$ be the corresponding diametral points of p, q respectively. Then the pairs of points $\bar{p}$, q; $\bar{p}$, $\bar{q}$; $\bar{q}$, p are non-diametral and hence they determine three arcs. The set of points $C_r = \text{arc } pq + \text{arc } \bar{p}q + \text{arc } \bar{p}\bar{q} + \text{arc } p\bar{q}$ is called a *great circle of S_r.* A great circle of S_r is evidently determined by two non-diametral points of S_r.

We may now prove the following theorem:

THEOREM 2. 2. *A great circle C_r of S_r is congruent to the $S_{1,r}$.*

Proof. Let p, q be two non-diametral points that determine the great circle C_r. Since $pq < \pi r$ we have p, q, $\bar{p}$, $\bar{q} \sim p'$, q', $\bar{p}'$, $\bar{q}'$, where the

† A point r lies *between* two points p, q provided $p \neq r$, $r \neq q$, $pr + rq = pq$. The triple is said to be *linear* in the order prq.

"primed" points are points of $S_{1,r}$. We have arc $pq \approx$ arc $p'q'$, arc $\bar{p}q \approx$ arc $\bar{p}'q'$, arc $\bar{p}\bar{q} \approx$ arc $\bar{p}'\bar{q}'$, arc $p\bar{q} \approx$ arc $p'\bar{q}'$. If now, x is any point of C_r then x is a point of one of the four arcs forming C_r. Let x' of $S_{1,r}$ be the point corresponding to x by virtue of the proper one of the above congruences. This gives a mapping of C_r upon $S_{1,r}$. To show this mapping is a congruent one let $x' \sim x$, $y' \sim y$. We wish to show that $xy = x'y'$.

We observe, first, that if x and y are both points of one of the four arcs, then $xy = x'y'$ by one of the above congruences of arcs. Suppose that x, y are not both points of the same arc and, to fix the ideas, let x be an interior point † of the arc pq and y an interior point of arc $\bar{p}q$. We have $qy + y\bar{p} = q\bar{p}$, $pq + q\bar{p} = p\bar{p}$, $p\bar{p} = py + y\bar{p}$. Adding these equalities, we obtain $pq + qy = py$. Since $p, x, q \approx p', x', q'$ of $S_{1,r}$, by Postulate D the four points p, q, x, y are congruent to four points p'', q'', x'', y'' of $S_{2,r}$, and since p, q, y are linear, the four points are on $S_{1,r}$. We may make a congruent transformation of $S_{1,r}$ into itself so that $p'' \to p'$, $q'' \to q'$, and $x'' \to x'$. This transformation sends y'' into a uniquely determined point y^*, and we have $p, q, x, y \approx p', q', x', y^*$.

Now, $q'y' + y'\bar{p}' = q'\bar{p}'$, $p'q' + q'\bar{p}' = p'\bar{p}'$, $p'\bar{p}' = p'y' + y'\bar{p}'$; adding, we obtain $p'q' + q'y' = p'y'$. From the above congruence and the preceding, $p'q' + q'y^* = p'y^*$, while $q'y' = qy = q'y^*$. Hence $p'y' = p'y^*$ and $y^* = y'$. Then $p, q, x, y \approx p', q', x', y'$ and $xy = x'y'$.

A space which is such that every quadruple of its points is congruent with some four points of the three-dimensional spherical space $S_{3,r}$ is said to have the spherical four-point property.

THEOREM 2. 3. *If the space S_r is convex, complete and diameterized, the space has the spherical four-point property.*

Proof. Let p_1, p_2, p_3, p_4 be four points of S_r. If the points contain a triple congruent with three points of $S_{1,r}$, then, by Postulate D, the four points are congruent with four points of $S_{2,r}$, and hence, in this case, the theorem is proved.

We suppose, now, that no triple contained in the four points is congruent with a triple of $S_{1,r}$. By Postulates A, B, C, $p_1, p_3, p_4 \approx p'_1, p'_3, p'_4$ of $S_{2,r}$, and $p_2, p_3, p_4 \approx p''_2, p''_3, p''_4$ of $S_{2,r}$. Since $p'_3p'_4 = p_3p_4 = p''_3p''_4$, we may transform $S_{2,r}$ congruently into itself so that $p''_3 \to p'_3$, $p''_4 \to p'_4$. The point p_2 is transformed into one of two points p_2^I, p_2^{II} which are images of each other in the great circle of $S_{2,r}$ determined by p'_3, p'_4. Denote this circle by $S_{1,r}(p'_3, p'_4)$. We have $p_2, p_3, p_4 \approx p_2^I, p'_3, p'_4 \approx p_2^{II}, p'_3, p'_4$. Now p'_1 is not on $S_{1,r}(p'_3, p'_4)$, so we may assume the labelling so that p'_1, p_2^I are both

† If x, y are not interior points of their respective arcs, the procedure is simplified.

on one of the two hemi-spheres determined by $S_{1,r}(p'_3,\ p'_4)$. Then $p'_1p_2^I < p'_1p_2^{II}$.

Let x' denote the intersection of the arc $p'_1p_2^{II}$ with $S_{1,r}(p'_3,\ p'_4)$†. The great circle $C_r(p_3,\ p_4)$ of S_r determined by $p_3,\ p_4$ is congruent to $S_{1,r}(p'_3,\ p'_4)$. In this congruence, let x of S_r correspond to x'. By Postulate D, $p_1, p_3, p_4, x \approx \tilde{p}_1, \tilde{p}_3, \tilde{p}_4, \tilde{x}$ of $S_{2,r}$. Since $\tilde{p}_1, \tilde{p}_3, \tilde{p}_4 \approx p_1, p_3, p_4 \approx p'_1, p'_3, p'_4$ of $S_{2,r}$, *not* $S_{1,r}$, a congruent transformation of $S_{2,r}$ into itself gives $p_1,\ p_3,\ p_4,\ x \approx p'_1,\ p'_3,\ p'_4,\ x^*$, with x^* uniquely determined. But $p'_3, p'_4, x' \approx p_3, p_4, x \approx p'_3, p'_4, x^*$, and p'_3, p'_4 are not diametral. Hence $x^* = x'$ and we have $p_1, p_3, p_4, x \approx p'_1, p'_3, p'_4, x'$. In a similar manner, we obtain $p_2, p_3, p_4, x \approx p_2^{II}, p_3, p_4, x'$. Now $p_1p_2 \leq p_1x + xp_2 = p'_1x' + x'p_2^{II} = p'_1p_2^{II}$.

Let the arc $p'_1\,p_2^I$, extended, meet $S_{1,r}(p'_3,\ p'_4)$ in a point y'. Let y of S_r correspond to y' in the congruence $C_r(p_3,\ p_4) \approx S_{1,r}(p'_3,\ p'_4)$. As above, we obtain $p_1,\ p_3,\ p_4,\ y \approx p'_1,\ p'_3,\ p'_4,\ y'$, and $p_2,\ p_3,\ p_4,\ y \approx p_2^I,\ p'_3,\ p'_4,\ y'$. Now $p_1p_2 \geq |\,p_1y - p_2y\,| = |\,p'_1y' - p_2^Iy'\,| = p'_1p_2^I$. Hence $p'_1p_2^I \leq p_1p_2 \leq p'_1p_2^{II}$.

Rotate the $S_{3,r}$ containing the $S_{2,r}$ half a revolution about $S_{1,r}(p'_3,\ p'_4)$ as axis. The point p_2^I is carried by this rotation into the point p_2^{II}. Let p'_x denote any point of its path. Then the function p'_1p_x is a continuous function which takes on the values $p'_1p_2^I$ and $p'_1p_2^{II}$, and consequently all values between these two. But p_1p_2 has been shown to lie between $p'_1p_2^I$ and $p'_1p_2^{II}$. Hence, the $S_{3,r}$ contains a point, say p'_2, such that $p'_1p'_2 = p_1p_2$, and p'_2 has the same distances from p'_3, p'_4 as p_2^I or p_2^{II}. Hence $p_1, p_2, p_3, p_4 \approx p'_1, p'_2, p'_3, p'_4$ of $S_{3,r}$, and the theorem is proved.‡

THEOREM 2.4. *Let C_r^1, C_r^2 be two distinct great circles of S_r having a point p in common. Then the set of points $C_r^1 + C_r^2$ can be isometrically imbedded in $S_{2,r}$.*

Proof. Let $p_1,\ p_2$ be points of $C_r^1,\ C_r^2$ respectively, distinct from p and not diametral to p. By Postulates B, C we have $p,\ p_1,\ p_2 \approx p',\ p'_1,\ p'_2$ of $S_{2,r}$. Let $S^1{}_{1,r},\ S^2{}_{1,r}$ denote the great circles of $S_{2,r}$ determined by the pairs $p',\ p'_1;\ p',\ p'_2$, respectively. By the preceding theorem, $C_r^1 \approx S^1{}_{1,r}$, $C_r^2 \approx S^2{}_{1,r}$. These two congruences give a mapping of the set $C_r^1 + C_r^2$ upon the set $S^1{}_{1,r} + S^2{}_{1,r}$. We must show this mapping is a congruent one.

Let $x,\ y$ be any two points of $C_r^1 + C_r^2$ and suppose them mapped into $x',\ y'$ respectively, by means of the above mapping. We show that $xy = x'y'$.

---

† If p'_1, p_2^{II} are diametral, any point on $S_{1,r}\ (p'_3, p'_4)$ may be taken as the point x'.

‡ In the proof given above of Theorem 2.3 the assumption that the space S_r is diameterized is needed to insure the existence in S_r of the great circle $C_r(p_3, p_4)$, and hence of the points x, y corresponding to x', y', respectively, of $S_{1,r}(p'_3, p'_4)$. A slightly different procedure, however, found by the writer while this paper was in press, establishes the theorem *without this assumption.* This applies also to Theorem 3.3 of the following section.

If x, y are both points of one of the two circles in S_r, then $xy = x'y'$ by virtue of one of the above congruences. Suppose that x is a point of ${C_r}^1$ and y a point of ${C_r}^2$. By Postulate D we have $p, x, p_1, p_2 \approx p'', x'', {p'}_1, {p''}_2$ of $S_{2,r}$. Now $p', {p'}_1, {p'}_2 \approx p, p_1, p_2 \approx p'', {p''}_1, {p''}_2$, and hence we may transform $S_{2,r}$ congruently into itself so that $p'' \to p'$, ${p''}_1 \to {p'}_1$, ${p''}_2 \to {p'}_2$. Since p, p_1, p_2 are evidently not on a great circle of S_r, the points p', ${p'}_1$, ${p'}_2$ are not on a great circle of $S_{2,r}$ and consequently the above congruent transformation sends x into a uniquely determined point x^*. We have $p, x, p_1, p_2 \approx p', x^*, {p'}_1, {p'}_2$ in $S_{2,r}$. Now, $p', x', {p'}_1 \approx p, x, p_1 \approx p', x^*, {p'}_1$ on ${S^1}_{1,r}$. Hence $x^* = x'$, and $p, x, p_1, p_2 \approx p', x', {p'}_1, {p'}_2$.

Consider, now, the quadruple p, p_2, x, y. We obtain by the usual procedure, $p, p_2, x, y \approx p', {p'}_2, x', y^*$ on $S_{2,r}$. As above, $p', {p'}_2, y' \approx p, p_2, y \approx p', {p'}_2, y^*$ on ${S^2}_{1,r}$. Hence $y^* = y'$ and $p, p_2, x, y \approx p', {p'}_2, x', y'$. Then $xy = x'y'$.

Using the result of Theorem 2. 3 it is readily shown that if three great circles of S_r have a point in common, their sum may be imbedded congruently in $S_{3,r}$.

THEOREM 2. 5. *Let p_1, p_2, p_3 be three points of the convex, complete space S_r which are not on a great circle of S_r. If we denote the three arcs they determine by A_{12}, A_{23}, A_{13}, then the set $A_{12} + A_{23} + A_{13}$ can be isometrically imbedded in $S_{2,r}$.*

The proof of this theorem is immediate. We remark that if the great circles ${C_r}^{12}$, ${C_r}^{23}$, ${C_r}^{13}$ exist in S_r, their sum may be imbedded in $S_{2,r}$. Since in Theorem 2. 5 we are not assuming that S_r is diameterized, these great circles may not exist.

3. We define n-dimensional sub-spaces $\Omega_{n,r}$ of S_r, assumed convex, complete, and diameterized, by recursion. Call a single point a zero-dimensional sub-space $\Omega_{0,r}$ of S_r and consider $n + 1$ points $p_0, p_1, \cdots, p_n$ of S_r which are congruent to $n + 1$ points of $S_{n,r}$ *not* of $S_{n-1,r}$. (If $n = 1$, we assume that the two points are not diametral). If the points $p_1, p_2, \cdots, p_n$ determine an $(n-1)$-dimensional sub-space $\Omega_{n-1,r}$ of S_r, the union of the great circles $\{p_0, x\}$, as x ranges over $\Omega_{n-1,r}$ is defined as an n-dimensional sub-space $\Omega_{n,r}$ of S_r with base points $p_0, p_1, \cdots, p_n$. Thus, for $n = 1$, $\Omega_{1,r}$ is the great circle of S_r determined by p_0, p_1 which has been shown to exist, is unique, and is congruent with the great circle $S_{1,r}$ of $S_{n,r}$ determined by ${p'}_0$, ${p'}_1$ where $p_0, p_1 \approx {p'}_0, {p'}_1$. It is obvious that no point of $\Omega_{n-1,r}$ is diametral with p_0.

This definition assigns a special rôle to the point p_0. It will be shown, however, that the base points uniquely determine the n-dimensional sub-space which is, therefore, independent of the particular vertex "isolated."

THEOREM 3.1. *Let p_0, p_1, p_2 be three points of the convex, complete diameterized space S_r, congruent with p'_0, p'_1, p'_2, three points of $S_{2,r}$, not of $S_{1,r}$. Then these three points determine exactly one two-dimensional sub-space $\Omega_{2,r}$ of S_r, with base points p_0, p_1, p_2, and this two dimensional sub-space is congruent with the two-dimensional spherical sub-space $S_{2,r}$ of $S_{n,r}$ determined by the three points p'_0, p'_1, p'_2. Moreover, the congruence of $\Omega_{2,r}$ with $S_{2,r}$ containing the congruence $p_0, p_1, p_2 \approx p'_0, p'_1, p'_2$ is unique.*

Proof. Let x be a point of $\Omega_{2,r}$. Then the great circle $C_r(p_1, p_2)$ of S_r contains a point σ such that x is a point of the great circle $C_r(p_0, \sigma)$ of S_r. By Theorem 2.2, $C_r(p_0, \sigma) \approx S_{1,r}(p'_0, \sigma')$.† Let x' of $S_{2,r}$ correspond to x by means of this congruence. This gives a mapping of $\Omega_{2,r}$ upon a sub-set of $S_{2,r}$.

If y is a point of $\Omega_{2,r}$ distinct from x, let y' of $S_{2,r}$ correspond to y by the above mapping. If y is contained in $C_r(p_0, \sigma)$ then $xy = x'y'$ by the congruence of great circles. If y is not a point of $C_r(p_0, \sigma)$ then there exists a point τ of $C_r(p_1, p_2)$ distinct from σ and $\bar{\sigma}$ such that y belongs to $C_r(p_0, \tau)$. Then p_0, σ, τ are three points of S_r which are not on a great circle of S_r and hence, by the remark following Theorem 2.5 the sum of the three great circles of S_r they determine may be isometrically embedded in $S_{2,r}$. This yields $xy = x'y'$ and the mapping is a congruent one.

Denote by $\Omega'_{2,r}$ the subset of $S_{2,r}$ into which $\Omega_{2,r}$ has been congruently transformed and let x' be any point of $S_{2,r}$. Now $p_0, p_1, p_2 \approx p'_0, p'_1, p'_2$ of $S_{2,r}$ not $S_{1,r}$, and x' is a point of a great circle of $S_{2,r}$ that contains p'_0 and a point s' of the great circle $S_{1,r}(p'_1, p'_2)$. Let s in S_r correspond to s' by virtue of the congruence $C_r(p_1, p_2) \approx S_{1,r}(p'_1, p'_2)$. Let x of S_r be selected on the great circle $C_r(p_0, s)$ so that $p_0, s, x \approx p'_0, s', x'$. Then x is a point of $\Omega_{2,r}$. Now since $\Omega_{2,r} \approx \Omega'_{2,r}$ there is a point x^* of $\Omega'_{2,r}$ that corresponds to x of $\Omega_{2,r}$ in this congruence. Evidently $x^* = x'$. Then x' is a point of $\Omega'_{2,r}$ and $\Omega'_{2,r} = S_{2,r}$.

To show that there is but a single two-dimensional sub-space with the points p_0, p_1, p_2 as base points, denote by $\tilde{\Omega}_{2,r}$ any other such space formed from these points. We have (1) $\Omega_{2,r} \approx S_{2,r}$ and (2) $\tilde{\Omega}_{2,r} \approx S_{2,r}$. Suppose x' is any point of $S_{2,r}$, distinct from p'_0, p'_1, p'_2 and let x of $\Omega_{2,r}$ $\tilde{x}$ of $\tilde{\Omega}_{2,r}$ be corresponding points by means of congruences (1), (2) respectively. Let x' be on the arc $p'_0 s'$, where s' is a point of $S_{1,r}(p'_1, p'_2)$ and let s of $\Omega_{2,r}$, $\tilde{s}$ of $\tilde{\Omega}_{2,r}$ correspond to s' in the appropriate congruences. Then

$$p_1, p_2, s \approx p'_1, p'_2, s' \approx p_1, p_2, \tilde{s}$$

† We have $C_r(p_1, p_2) \approx S_{1,r}(p'_1, p'_2)$ and $\sigma' \sim \sigma$ by this congruence.

where $s, \bar{s}$ are on $C_r(p_1, p_2)$. Then $s = \bar{s}$, and it follows immediately that $x = \bar{x}$, for

$$p_0, s, x \approx p'_0, s', x' \approx p_0, \bar{s}, \bar{x} \approx p_0, s, \bar{x}.$$

Finally, we observe that the congruence $\Omega_{2,r} \approx S_{2,r}$ containing the congruence $p_0, p_1, p_2 \approx p'_0, p'_1, p'_2$ is unique, for the contrary implies a congruent transformation of $S_{2,r}$ into itself leaving the points p_0, p_1, p_2 fixed and which is not an identical transformation. This, of course, is impossible.

The analogous theorem concerning the n-dimensional sub-space is proved by complete induction. The procedure employed in the proof of Theorem 3. 1 is easily generalized and the following theorem is obtained.

THEOREM 3. 2. *Let p_i $(i = 0, 1, \cdots, n)$ be $n+1$ points of the complete, complex, diameterized space S_r which are congruent with $n+1$ points p'_i $(i = 0, 1, \cdots, n)$ of $S_{n,r}$, not of $S_{n-1,r}$. There is exactly one n-dimensional sub-space $\Omega_{n,r}$ of S_r that has the points p_i $(i = 0, 1, \cdots n)$ as base points and the congruence $p_0, p_1, \cdots, p_n \approx p'_0, p'_1, \cdots, p'_n$ induces a unique congruence of $\Omega_{n,r}$ with the $S_{n,r}$.*

The existence of the sub-space $\Omega_{n,r}$ congruent with $S_{n,r}$ depends on the space S_r containing $n+1$ points that may be imbedded isometrically in the $S_{n,r}$. By Theorem 2. 3, every four points of S_r, assumed convex, complete and diameterized, is congruent to four points of $S_{3,r}$; that is S_r has the spherical four-point property. We show, by complete induction that this implies that such a space has the spherical n-point property for every integer n.

THEOREM 3. 3. *The convex, complete, diameterized space S_r has the spherical n-point property for every integer n.*

Proof. We suppose that S_r has the spherical k-point property for every $k \leq n$, where n is a fixed integer greater than three. Let $p_0, p_1, \cdots, p_n$ be any $n+1$ points of S_r. By hypothesis, any k of these points, $k \leq n$, are congruent with k points of $S_{k-1,r}$. Now either the given $n+1$ points are congruent with $n+1$ points of $S_{n-1,r}$ (in which case the theorem is proved) or n points of the set may be selected to form the base points of an $(n-1)$-dimensional subspace $\Omega_{n-1,r}$ of S_r which does not contain the remaining point.

In this case we assume the labelling so that

$$p_0, p_2, \cdots, p_n \approx p'_0, p'_2, \cdots, p'_n \text{ of } S_{n-1,r},$$

not of $S_{n-2,r}$, and

$$p_1, p_2, \cdots, p_n \approx \bar{p}_1, \bar{p}_2, \cdots, \bar{p}_n \text{ of } S_{n-1,r},$$

since the spherical k-point property is assumed for $k \leq n$. We may transform $S_{n-1,r}$ congruently into itself so that $\bar{p}_2 \to p'_2, \cdots, \bar{p}_n \to p'_n$. This transforms

$\bar{p}_1$ into a point p_1^*, which has its distances from the $n-1$ points $p'_2, \cdots, p'_n$ determined. Since the points $p'_2, p'_3, \cdots, p'_n$ are not on $S_{n-3,r}$ there are *exactly two distinct positions* for the point p_1^* on the $S_{n-1,r}$ containing $p_1^*, p'_2, \cdots, p'_n$. We denote the two positions of p_1^* by p_1^I, p_2^{II}; these two points are evidently images of each other in the $S_{n-2,r}$ determined by $p'_2, p'_3, \cdots, p'_n$. We have

$$p_1, p_2, \cdots, p_n \approx p_1^I, p'_2, \cdots, p'_n \approx p_1^{II}, p'_2, \cdots, p'_n.$$

Since p'_0 does not lie on the $S_{n-2,r}$ containing $p'_2, \cdots, p'_n$, $p'_0 p_1^I \neq p'_0 p_1^{II}$. Assume the labelling so that p'_0, p_1^I are *both* on one of the two "hemi-spheres" of $S_{n-1,r}$ determined by $S_{n-2,r}(p'_2, \cdots, p'_n)$. Then $p_0 p_1^I < p_0 p_1^{II}$.

The arc † $p'_0 p_1^{II}$ intersects $S_{n-2,r}(p'_2, \cdots, p'_n)$ in a point x'. The $(n-2)$-dimensional space $\Omega_{n-2,r}$ with base points $p_2, p_3, \cdots, p_n$ of S_r is congruent to $S_{n-2,r}(p'_2, p'_3, \cdots, p'_n)$. In this congruence let x of S_r correspond to x'. This congruence is contained in both of the congruences

(1) $$\Omega_{n-1,r}(p_0, p_2, \cdots, p_n) \approx S_{n-1,r}(p'_0, p'_2, \cdots, p'_n)$$

(2) $$\Omega_{n-1,r}(p_1, p_2, \cdots, p_n) \approx S_{n-1,r}(p_1^I, p'_2, \cdots, p'_n) \approx S_{n-1,r}(p_1^{II}, p'_2, \cdots, p'_n).$$

By congruence (1), $p_0 x = p'_0 x'$ and by congruence (2), $x p_1 = x' p_1^{II}$. Now, by the triangle inequality, $p_0 p_1 \leq p_0 x + x p_1 = p'_0 x' + x' p_1^{II} = p'_0 p_1^{II}$. Again, the arc $p'_0 p_1^I$ extended, intersects $S_{n-2,r}(p'_2, \cdots, p'_n)$ in a point y'. We have, as above, $p_0 y = p'_0 y'$, $p_1 y = p_1^I y'$ where y of S_r corresponds to y' by the congruence $\Omega_{n-2,r}(p_2, \cdots, p_n) \approx S_{n-2,r}(p'_2, \cdots, p'_n)$. Then

$$p_0 p_1 \geq | p_0 y - p_1 y | = | p'_0 y' - p_1^I y' | = p'_0 p_1^I.$$

Hence $$p'_0 p_1^I \leq p_0 p_1 \leq p'_0 p_1^{II}.$$

Suppose the $S_{n,r}$ containing the $S_{n-1,r}(p'_0, p_1^I, p_1^{II}, p'_2, \cdots, p'_n)$ is rotated 180 degrees about the $S_{n-2,r}(p'_2, \cdots, p'_n)$ as axis. This rotation keeps fixed the points $p'_2, \cdots, p'_n$, while the point p_1^I is carried over to the point p_1^{II}. Let p'_x denote any point of its path. Keeping the label of point p'_0 fixed, the function $p'_0 p'_x$ is a continuous function which assumes the values $p'_0 p_1^I$ and $p'_0 p_1^{II}$ and hence for any value between these two, such as $p_0 p_1$, there is a point, say p'_1, of $S_{n,r}$, such that $p'_0 p'_1 = p_0 p_1$. Hence

$$p_0, p_1, p_2, \cdots, p_n \approx p'_0, p'_1, p'_2, \cdots, p'_n \text{ of } S_{n,r}$$

and the theorem is proved.‡

4. Let 0 be a point of Hilbert space. The set $\{x\}$ of points of the

† If $p'_0 p_1^{II}$ are diametral, any point of $S_{n-2}(p'_2, \cdots, p'_n)$ may be taken as the point x'.

‡ If $p_0 p_1$ equals $p'_0 p_1^I$ or $p'_0 p_1^{II}$ the $n+1$ points are evidently congruent with $n+1$ points of $S_{n-1,r}$.

space whose distance from 0 is equal to a positive constant r is called the sphere with center 0 and radius r in Hilbert space, and is denoted by H_r.

Let p_i, p_j be any two points of S_r and denote by α_{ij} the angle p_ip_j/r radians. If $p_1, p_2, \cdots, p_k$ are any k points of S_r, we denote the axisymmetric determinant

$$\begin{vmatrix} 1 & \cos\alpha_{12} & \cdots & \cos\alpha_{1k} \\ \cos\alpha_{21} & 1 & \cdots & \cos\alpha_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \cos\alpha_{k1} & \cos\alpha_{k2} & \cdots & 1 \end{vmatrix}$$

by $\Delta_k(p_1, p_2, \cdots, p_k)$.

If the space S_r is such that there exists an integer k such that the determinant Δ_k formed for every set of k points of S_r vanishes, we say the space S_r is *indexed*. If S_r is indexed there is evidently a smallest integer for which S_r has this property. Call this integer the index of S_r.

We now prove the fundamental theorem characterizing metrically the n-dimensional spherical space $S_{n,r}$.

THEOREM 4.1. *The convex, complete, diameterized and indexed space S_r with index $n+2$ is congruent with $S_{n,r}$.*

Proof. Since S_r has index $n+2$ there exists at least one set of $n+1$ points of S_r for which Δ_{n+1} does not vanish, while $\Delta_{n+2}=0$ for every set of $n+2$ points of S_r. By Theorem 3.3, every set of $n+1$ points of S_r is congruent with $n+1$ points of $S_{n,r}$. Hence, since Δ_{n+2} vanishes, every set of $n+2$ points of S_r is congruent with $n+2$ points of $S_{n,r}$.†

Consider, now, any set of $n+3$ points of S_r. Since each set of $n+2$ points contained in these $n+3$ points may be imbedded isometrically in $S_{n,r}$, the $n+3$ points are either congruent with $n+3$ points of $S_{n,r}$ or they form a pseudo r-spheric $(n+3)$-tuple. But each set of $n+3$ points of S_r is, by Theorem 3.3, congruent with $n+3$ points of $S_{n+2,r}$; that is, $\Delta_{n+3}\geq 0$; while the determinant Δ_{n+3} formed for a pseudo r-spheric $(n+3)$-tuple is negative.‡ Hence no set of $n+3$ points of S_r can be pseudo r-spheric and therefore each $(n+3)$-tuple is congruent with an $(n+3)$-tuple of $S_{n,r}$. Since S_n has the congruence order $n+3$, this implies that S_r is congruent to a subset of $S_{n,r}$.

Now S_r contains a set of $n+1$ points for which Δ_{n+1} does not vanish. These points are, therefore, not congruent with $n+1$ points of $S_{n-1,r}$ and

† See Theorem II_n of Blumenthal and Garrett, "Characterization of spherical and pseudo-spherical sets of points," *American Journal of Mathematics*, Vol. 55 (1933), p. 632.

‡ *Loc. cit.*, *Theorems* I_n, IV_n, p. 632.

hence they determine a sub-space $\Omega_{n,r}$ of S_r which is congruent to $S_{n,r}$. Hence $S_r \approx S_{n,r}$ and the theorem is proved.

The metric characterization of the sphere H_r in Hilbert space is given by the following theorem.

THEOREM 4. 2. *If the convex, complete, diameterized, separable space S_r is not indexed it is congruent to H_r, the sphere in Hilbert space.*

Proof. By Theorem 3. 3, every k points of S_r is congruent with k points of $S_{k-1,r}$; that is, $\Delta_k \geq 0$ for every integer k. Since S_r is separable, we may conclude that S_r is congruent to a subset of H_r.†

Since S_r is not indexed, for every integer k, S_r contains at least one set of k points for which Δ_k does not vanish and hence is positive. Then S_r is not isometric with $S_{n,r}$ for any value of n.

It follows readily from the completeness of S_r that $S_r \approx H_r$.

If each of the $n+3$ sets of $n+2$ points contained in $n+3$ points of a semi-metric space is congruent with $n+2$ points of $S_{n,r}$, while the $n+3$ points are not congruent with $n+3$ points of $S_{n,r}$, the set is said to form a pseudo r-spheric $(n+3)$-tuple. From the foregoing theorems it follows that a convex, complete, diameterized space S_r cannot contain pseudo r-spheric sets. Thus, for example, the set of five points $p_1, p_2, \cdots, p_5$, with the distance of each pair defined to be $r \cdot \cos^{-1}(-1/3)$, has each of its quadruples isometric with four points of $S_{2,r}$, while the five points are clearly not isometric with five points of $S_{2,r}$. As a consequence of our theorems these five points cannot be imbedded in a space S_r which is convex, complete, and diameterized.

It may be observed that the theorems obtained by W. A. Wilson in his investigation of metric spaces in which each four points are congruent with four points of a euclidean space, are limiting cases of the theorems presented in this paper, while Postulate D is weaker than the four-point condition assumed by Wilson.‡

THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

---

† Menger has shown (*Ergebnisse Eines Mathematischen Kolloquiums*, Wien, Heft 1 (1931), p. 26) that a necessary and sufficient condition for a separable semi-metric space R to be congruent with a subset of Hilbert space is that for every integer k and for every k points $p_1, p_2, \cdots, p_k$ of R, $\operatorname{sign} D(p_1, p_2, \cdots, p_k) = \operatorname{sign}(-1)^k$ or 0, where $D(p_1, p_2, \cdots, p_k)$ is the well-known volume-determinant. We obtain the result above by replacing p_ip_j by $2r \cdot \sin p_ip_j/2r$ in Menger's theorem.

It may be pointed out that there is a typographical error in the statement of this theorem occurring in Menger, "Die Metrik des Hilbert'schen Raumes," *Wiener Akademischer Anzeiger*, 1928, No. 12, p. 160.

‡ W. A. Wilson, "A relation between metric and euclidean spaces," *American Journal of Mathematics*, Vol. 54 (1932), pp. 505-517.

ON CERTAIN TYPES OF CONTINUOUS TRANSFORMATIONS OF METRIC SPACES.*

By W. A. WILSON.

1. The literature on topology produced during the last two decades has by its great volume tended to divert attention from the study of geometrical properties which are not invariant under continuous transformations. It may well be that with more restricted transformations the methods used in topological problems can be applied with equal success to the study of non-topological properties of geometrical configurations. It is the purpose of this article to develop some of the characteristics of a certain class of continuous transformations and their application to geometry,—with the hope that they indicate the possibility of further progress in the direction mentioned above.

2. *Definition of spread at a point.* Let $X = \{x\}$ be a metric space and, as usual, let xx' denote the distance between the points x and x'. If $Y = \{y\}$ is another metric space which is the continuous image of X by a transformation $y = f(x)$, one-valued and defined over X, there are three possibilities at any limiting point a of X. First, there may be a constant $m_f(a)$ such that, if y and y' correspond to distinct points x and x', respectively, $\lim(yy'/xx') = m_f(a)$ as x and x' approach a. (It is understood that either x or x' may coincide with a.) In this case we say that the transformation has a *finite spread at* a and denote the spread by $m_f(a)$. Second, it may happen that $\lim(yy'/xx') = \infty$; in this case we say that the *spread is infinite at* a and write $m_f(a) = \infty$. Finally, yy'/xx' may approach no limit as x and x' approach a, and we then say that the *spread is indeterminate at* a.

That the last case can arise effectively is seen in a homeomorphism between a straight line and a broken line. Examples of the first two are easily invented. We shall be chiefly concerned with transformations for which the spread is defined at each point and is constant.

3. *Elementary properties of the spread.* From the definition it is apparent that the spread is somewhat analogous to the derivative. In fact, it is easy to show that, if x and y are real variables and $y = f(x)$ is continuous and differentiable, continuity of $f'(x)$ at $x = a$ implies that $m_f(a)$ is finite, and conversely; moreover, $m_f(a) = |f'(a)|$. As in the case of derivatives, the following properties are deducible from the definition:

* Presented to the Society, September, 1934.

I. *Let* $y = f(x)$ *and its inverse* $x = g(y)$ *define a homeomorphism between the metric spaces* X *and* Y. *If* $b = f(a)$ *and* $m_f(a)$ *exists, finite or infinite, so does* $m_g(b)$. *If* $m_f(a) \neq 0$ *and is finite,* $m_g(b) = 1/m_f(a)$; *if* $m_f(a) = 0$, $m_g(b) = \infty$; *and if* $m_f(a) = \infty$, $m_g(b) = 0$.

II. *Let* $X = \{x\}$ *be a compact metric space and the continuous transformation* $y = f(x)$ *have a finite spread at each limiting point of* X. *Then* yy'/xx' *is bounded in* X.

III. *Let* $X = \{x\}$ *be a compact metric space and the transformation* $y = f(x)$ *have a finite spread* $m_f(x)$ *which is continuous in* X. *Then* yy'/xx' *converges to* $m_f(x)$ *uniformly in* X.

The proof of I is like that of the corresponding theorem on derivatives. If we assume that II fails, it follows from the compactness of X and the continuity of $f(x)$ that there is some point a and points x and x' converging to a for which $yy'/xx' \to \infty$, in contradiction to the hypothesis that $m_f(a)$ is finite. If III fails, there is a sequence of points $\{a_i\}$ converging to a point a and points x_i and x'_i such that $a_i x_i \to 0$ and $a_i x'_i \to 0$, but $| m_f(a_i) - y_i y'_i / x_i x'_i |$ exceeds a preassigned positive ϵ. This contradicts the hypothesis that $m_f(a_i) \to m_f(a)$.

4. *Applications to rectifiability.* Let us suppose that $y = f(x)$ defines a homeomorphism between the metric spaces X and Y. If the spread is finite at each point, it is evident by Property II of the previous section and the usual definition of length of arc, that rectifiable arcs in X go over into rectifiable arcs in Y. If the spread is continuous and for some arc in X the limit of the ratio of the length of a sub-arc ax to the distance ax is any constant m, as x approaches a, this property is also preserved in the image by virtue of Property III. When the spread is a constant $k \neq 0$, the length of the image of an arc of length s is easily seen to be ks. (In this case, if intrinsic distances * are used, the transformation is one of similitude.) If the spread is everywhere infinite, it follows from Properties I and III that the image of no rectifiable arc is rectifiable. If the spread is everywhere zero, X can contain no rectifiable arc, since the image of such an arc would be a point.

On the other hand we can have a homeomorphic transformation with a constant zero spread of a metric space which contains no rectifiable arc, as shown by the following example. Let $Y = \{y\}$ be the linear interval $0 \leqq y \leqq 1$. Let $X = \{x\}$ be a metric space obtained from Y by the con-

* Compare the writer's paper, "On rectifiability in metric spaces," *Bulletin of the American Mathematical Society*, vol. 38, pp. 419-426.

tinuous transformation $x = g(y)$ such that, for any two points y, y' and the corresponding points x, x', we have $xx' = \sqrt{yy'}$. It is easily seen that this transformation is a homeomorphism and that X is a simple arc. If $y = f(x)$ is the inverse of $x = g(y)$, we see that $yy'/xx' = xx'$; hence $m_f(x) = 0$ at every point of X and $m_g(y) = \infty$ at every point of Y.

The arc X has some remarkable properties. Any $n+1$ points $a_0, a_1, \cdots, a_n$ may be arranged so that, if $b_i = f(a_i)$, $b_i < b_{i+1}$. Setting $r_i = b_{i-1}b_i$, we have $a_0a_1 = \sqrt{r_1}$, $a_0a_2 = \sqrt{r_1 + r_2}$, $a_0a_3 = \sqrt{r_1 + r_2 + r_3}$, $\cdots$, $a_1a_2 = \sqrt{r_2}$, etc. Substituting these values in the $(n+1)$-point determinant D_n used by Menger * as a criterion of imbeddability in Euclidean space, we find by easy algebra that $D_n = (-1)^{n+1} \cdot 2^n \cdot r_1 \cdot r_2 \cdot \cdots r_n$. Hence the arc can be imbedded † in Hilbert space, no $n+1$ of its points can be imbedded in an $(n-1)$-dimensional Euclidean space, and in particular no three of its points lie in a straight line. The curve is extremely crinkly; not only is no part of it rectifiable, but, if a, b, and c are any three points in order, abc is a right angle.

5. *Unrestricted regular transformations.* The foregoing example suggests the following definition. Let $v = \phi(u)$ be a one-valued real function defined for $u \geqq 0$ and subject to these conditions: (a) $\phi(u) > 0$ if $u > 0$; (b) $\phi(0) = 0$; (c) $\phi(u)$ is continuous; (d) if $u_1 + u_2 \geqq u_3$, then $\phi(u_1) + \phi(u_2) \geqq \phi(u_3)$. If $y = f(x)$ is a continuous transformation of a metric space X into a metric space Y such that, if u is the distance between two points in X, the distance between the image points in Y is $v = \phi(u)$, where $\phi(u)$ satisfies Conditions (a) to (d) above, we say that the transformation is an *unrestricted regular* transformation and call $\phi(u)$ the *scale function* of the transformation.

Such a transformation is regular in the sense that the spread is constant (as will be seen soon) and congruent figures go over into congruent figures. It is unrestricted in the sense that it can be applied to any metric space. A transformation whose scale function satisfies the given conditions except for (d) would preserve congruences and might be applicable to a certain space X, but not applicable to a metric space containing X. Simple examples of functions satisfying the given requirements are $v = \sqrt{u}$, $v = \log(1+u)$, and $v = \sin u$ for $u \leqq \pi/2$ and $v = 1$ for $u > \pi/2$. The function $v = u^2$ does not satisfy Condition (d) and so a transformation involving this relation

* "Untersuchungen über allgemeine Metrik," *Mathematische Annalen*, vol. 100, pp. 119, 120.

† I. e., is congruent with a sub-set of Hilbert space.

between distances would be impossible unless for every triple of points in the given space $|(ab)^2 - (ac)^2| \leqq (bc)^2 \leqq (ab)^2 + (ac)^2$.

Any scale function $v = \phi(u)$ of an unrestricted regular transformation has a variety of properties,* of which the following are important:

I. It is monotone increasing.

II. $\phi'(0)$ exists, finite or infinite, and $\phi'(0) > 0$.

III. For any value of u, $\limsup_{\Delta u \to 0} (\Delta\phi/\Delta u) \leqq \phi'(0)$.

IV. If the inverse, $u = \psi(v)$, is also a scale function of an unrestricted regular transformation, v/u is a constant.

Property I is easily deduced from Conditions (a) to (d). To prove Property II let M be the upper bound of $\phi(u)/u$; then $\limsup_{u \to 0} [\phi(u)/u] \leqq M$. If M is finite, there is, for every $\epsilon > 0$, some value $\bar{u}$ of u for which $\phi(\bar{u})/\bar{u} > M - \epsilon/2$. Now $\phi(u)/u$ is continuous; hence for some $\delta > 0$ and any u such that $|u - \bar{u}| < \delta$, $\phi(u)/u > M - \epsilon$. If $u' < \delta$, there is an integer n for which $|nu' - \bar{u}| < \delta$. Moreover, by repeated application of Condition (d),

$$\phi(u')/u' \geqq \phi(nu')/nu' > M - \epsilon.$$

Hence $\liminf_{u \to 0} [\phi(u)/u] \geqq M$. A similar argument applies if $M = \infty$. Therefore $\phi'(0)$ exists and equals M. Clearly $M \neq 0$, since $\phi(u) > 0$ if $u > 0$. Property III follows from Condition (d) in the above definition, and Property IV is a consequence of III and the fact that $\phi'(0) \cdot \psi'(0) = 1$.

THEOREM. *Let $X = \{x\}$ be a dense metric space and $y = f(x)$ be an unrestricted regular transformation whose scale function is $v = \phi(u)$. Then $y = f(x)$ defines a homeomorphism and its spread is $\phi'(0)$ at each point of X.*

Proof. The inverse transformation $x = g(y)$ is one-valued, since $\phi(u) > 0$ for $u > 0$. If $g(y)$ is not everywhere continuous, there is a point y_0 and a sequence $\{y_i\}$ converging to y_0, such that, if $x_0 = g(y_0)$ and $x_i = g(y_i)$, $x_i x_0$ exceeds some positive number k for all values of i. If we let $u_i = x_i x_0$, we have $v_i = \phi(u_i) \geqq \phi(k) > 0$. This is impossible, as $v_i = y_i y_0$ and $y_i \to y_0$. Hence $g(y)$ is both one-valued and continuous, and the transformation is a homeomorphism.

If a is any point of X, there are points x and x' in any vicinity of a. If we set $u = xx'$, $yy'/xx' = v/u$ and, as x and x' approach a, $u \to 0$. Hence $m_f(a) = \lim_{u \to 0} (v/u) = \phi'(0)$.

* Compare also F. Toranzos, "Über die Funktionel-Ungleichung $f(x) + f(y)$ $f(x + y)$," *Rev. mat. hisp.-amer.*, vol. 2, pp. 109-113.

6. The theorem just proved shows that the discussion of § 4 is valid for the class of unrestricted regular transformations. Furthermore, if the inverse of such a transformation is also regular and unrestricted, Property IV of scale functions shows that the transformation is one of similitude.

This is not, however, universally the case. The functions $v = \sqrt{u}$ and $v = \log(1 + u)$, mentioned earlier, are instances in point. Such transformations, when applied to ordinary spaces, produce interesting new spaces. For example, if $\phi'(0) = k \neq \infty$ and the given space is complete and convex, the image is complete and each pair of points is joined by an arc of length k times the distance between the corresponding points in the given space, but the image is not convex unless the transformation is one of similitude. If we take as the space X a unit square in the plane and apply a transformation with the scale function $v = \sqrt{u}$ (here $\phi'(0) = \infty$), the image Y is a simple two-cell, the images of segments in X of equal length being congruent simple arcs of the type described in § 4. It would be interesting to know whether this and similar configurations can be imbedded in Hilbert space.

7. *Angles and conformality.* Let A, B, and C be points in a metric space, $a = BC$, $b = AC$, and $c = AB$. Let us call the symbol BAC an *angle* with vertex A and define its value by the formula $\cos BAC = (b^2 + c^2 - a^2)/2bc$, taking the value between 0 and π. This definition is possible by virtue of the metric triangle inequality. If ρ and σ are two rays * with the common initial point A, and B and C are other points on ρ and σ, respectively, we say that ρ and σ make an angle (ρ, σ) if $\lim BAC$ exists when B and C approach A. A homeomorphic transformation in which angles between rays are preserved will be called *conformal.*

In general metric spaces angles defined in this manner obviously lack many important properties usually associated with angles and a further investigation of the types of spaces admitting these properties and of conditions for the existence of angles between rays is needed. We can, however, obtain one interesting result without restricting our space.

THEOREM. *Let X and Y be homeomorphic metric spaces and the transformation of X into Y be an unrestricted regular transformation whose scale function $\phi(u)$ has a derivative continuous at $u = 0$. Then the transformation is conformal.*

Proof. Let A be a fixed point, and B and C be variable points in X. Let $a = BC$, $b = AC$, and $c = AB$. Let the corresponding points and distances in Y be denoted by primed letters. To prove our theorem we must

* I. e., topological images of half-lines.

show that, if $\lim \cos BAC$ exists as B and C approach A, then $\lim \cos B'A'C'$ exists and is the same.

For any sequence of positions of B and C converging to A at least one of the following three cases for some sub-sequence will occur: (I) $a/b \to \infty$ or $a/c \to \infty$; (II) $a/b \to 0$ or $a/c \to 0$; (III) $a/b, a/c, b/a,$ and c/a are bounded.

Case I. Let $a/b \to \infty$. By algebra

$$1 - \cos B'A'C' = (1/2)[1 + (a' - b')/c'] \cdot [1 + (a' - c')/b'].$$

By the mean value theorem we have $a' - b' = (a - b) \cdot \phi'(u')$ and $c' = c \cdot \phi'(u'')$, where $a < u' < b$ and $0 < u'' < c$. We can then write

$$(a' - b')/c' = [(a - b)/c] \cdot [\phi'(u')/\phi'(u'')].$$

As a, b, and c approach zero, $\lim [\phi'(u')/\phi'(u'')] = 1$. Since $a/b \to \infty$, it follows by the metric triangle inequality that $\lim [(a - b)/c] = 1$. Hence we know that $\lim [1 + (a' - b')/c'] = \lim [1 + (a - b)/c] = 2$. In consequence of the existence of $\lim (1 - \cos BAC)$ and of $\lim [1 + (a - b)/c]$, it follows that $\lim [(a - c)/b]$ exists. By an argument like that above, we then have $\lim [1 + (a' - c')/b'] = \lim [1 + (a - c)/b]$. Consequently, $\lim (1 - \cos B'A'C') = \lim (1 - \cos BAC)$. If $a/c \to \infty$, we proceed in exactly the same manner.

Case II. Let $a/b \to 0$ or $a/c \to 0$. The procedure is similar to that in Case I, with the exception that $\lim [(a - b)/c] = -1$, or $\lim [(a - c)/b] = -1$.

Case III. If a/c, a/b, b/a, and c/a are bounded, it is shown by the metric triangle inequality that b/c and c/b are also bounded. By simple algebra we find that

$$2(\cos B'A'C' - \cos BAC)$$
$$= (b/c)[b'c/bc' - 1] + (c/b)[bc'/b'c - 1] + (a^2/bc)[1 - bca'^2/b'c'a^2].$$

Since $\phi'(0)$ is finite and not zero, each bracket converges to zero. On the other hand, the coefficients are all bounded. Hence the above equation gives $\lim \cos B'A'C' = \lim \cos BAC$.

Thus we have shown that every sequence of points B' and C' which are images of points B and C converging to A contains a sub-sequence for which $\lim \cos B'A'C' = \lim \cos BAC$. But it then follows from the theory of limits that the same equation is true for all sequences. Hence the theorem is proved.

8. *Smoothness of arcs.* Let A, B, and C be points of an oriented simple arc, with B and C on the right of A and the letters so assigned that $AB \leqq AC$. If $\lim [(AB + BC - AC)/AB] = 0$ as B and C approach A, we say that the arc is *smooth on the right* at A. Similarly we define smoothness on the left.

If the arc is smooth on both sides of A and also for points B and C approaching A from opposite sides, with $AB \leqq AC$, $\lim [(AB + AC - BC)/AB] = 0$, we say that the arc is *smooth about* A. In another paper by the writer * it is shown that for arcs in a Euclidean space these intrinsic conditions are necessary and sufficient for the existence of unilateral tangents and of tangents.

If the points are in the order A, B, C and both $\lim [(AB + BC - AC)/AB] = 0$ and $\lim [(AB + BC - AC)/BC] = 0$, we say that the arc is *strongly smooth on the right* at A. If the order is C, B, A, we have strong smoothness in the left. By application of the metric triangle inequality we can show that these conditions imply the condition for unilateral smoothness above. If an arc is strongly smooth on both sides of A and also smooth about A, we call it *strongly smooth about* A. It is readily seen from the plane cosine law that the conditions for unilateral strong smoothness are equivalent to requiring that angle $BAC \to 0$ and angle $ABC \to \pi$ as B and C approach A. If the arc lies in a Euclidean plane and has a tangent at each point B of a sub-arc AC, it then follows that strong smoothness on the right is equivalent to the requirement that the tangent turns in such a manner as to approach the right-hand tangent at A as B approaches A. Thus strong smoothness is a generalization for abstract spaces of the idea of a continuously turning tangent.

THEOREM. *Let X and Y be homeomorphic metric spaces and the transformation of X into Y be unrestricted and regular and its scale function $\phi(u)$ have a derivative continuous at $u = 0$. If a simple arc in X has smoothness of any kind at a point A, its image in Y has the same kind of smoothness at the corresponding point A'.*

Proof. Let us take the case that the given arc is smooth on the right at A; the other cases are proved in exactly the same way. Let B and C be points of the arc at the right of A, $a = BC$, $b = AC$, $c = AB$, and $c \leqq b$. Let the corresponding points and distances in Y be denoted by primed letters.

By definition $\lim [(c + a - b)/c] = 0$ as B and C approach A, whence $\lim [(a - b)/c] = -1$. By the mean value theorem $a' - b' = (a - b) \cdot \phi'(u')$ and $c' = c \cdot \phi'(u'')$, where $a < u' < b$ and $0 < u'' < c$. This gives $(a' - b')/c' = [(a - b)/c] \cdot [\phi'(u')/\phi'(u'')]$. By the continuity of $\phi'(u)$, $\lim [\phi'(u')/\phi'(u'')] = 1$. Hence $\lim [(a' - b')/c'] = \lim [(a - b)/c] = -1$, and $\lim [(c' + a' - b')/c'] = 0$, which was to be proved.

YALE UNIVERSITY,
NEW HAVEN, CONN.

* "On angles in certain metric spaces," *Bulletin of the American Mathematical Society*, vol. 38, pp. 580-588.

DISCONTINUOUS SOLUTIONS IN THE NON-PARAMETRIC PROBLEM OF MAYER IN THE CALCULUS OF VARIATIONS.*

By WILLIAM T. REID.

1. Introduction. By a discontinuous solution in a calculus of variations problem is meant an extremizing arc having one or more corners; that is, the derivatives of the functions defining the arc have one or more points of discontinuity. The treatment of discontinuous solutions has been largely limited to problems phrased in parametric form [see IV and V; the reader is referred to V for references to earlier literature on the subject].† In fact, Graves [V, p. 834] has stated that "it seems to be necessary to cast the problem in parametric form to get an effective extension of the Jacobi condition by means of the second variation."

The present paper treats discontinuous solutions for the general problem of Mayer in non-parametric form, and in terms of the characteristic numbers of a boundary value problem associated with the second variation there is given an effective extension of the Jacobi condition. Sufficient conditions are then established by the method used in the case of extremal arcs by Bliss and Hestenes and by Hestenes [VIII and IX]. The reader is also referred to the bibliographies at the end of these papers for references to earlier literature on the problem of Mayer. For the accessory boundary value problem in terms of which the analogue of the Jacobi condition is phrased the boundary conditions apply at more than two points, but it is significant that this problem may be transformed into one for which the boundary conditions apply at just two points, and of the type recently considered by the author [VII; also X, § 6]. In view of a recent paper by the author [X], it follows that an analogue of the Jacobi condition may equally well be phrased in a manner which is a direct generalization of that given by Bliss for the problem of Bolza, and also used by Hestenes for the general problem of Mayer [IX, p. 483].

In §§ 2 and 4 there are given conditions satisfied by a minimizing arc for the problem here considered, while § 3 considers the construction of families of extremaloids. The second variation, an analogue of Jacobi's condition in terms of an associated boundary value problem, together with the definition and determination of conjugate points, are discussed in §§ 5-8. In § 9 there is proved an auxiliary theorem which is fundamental in the

* Presented to the American Mathematical Society, June 23, 1933.

† Roman numerals in brackets refer to the bibliography at the end of this paper.

construction of a field of extremaloids as defined in § 10, and as used in § 11 to establish sufficient conditions for a strong relative minimum.

2. The first necessary condition for a minimum, and preliminary remarks. The general problem of Mayer as formulated by Bliss [I] is that of finding in a class of arcs

$$y_i = y_i(x) \qquad (i = 1, \cdots, n;\ x_1 \leq x \leq x_2) \tag{2.1}$$

satisfying a system of differential equations and end-conditions

$$\phi_\alpha[x, y, y'] = 0 \qquad (\alpha = 1, \cdots, m < n) \tag{2.2}$$

$$\psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0 \qquad (\mu = 1, \cdots, p \leq 2n + 1) \tag{2.3}$$

one which minimizes a function of the form

$$g[x_1, y(x_1), x_2, y(x_2)]. \tag{2.4}$$

As is customary, we concentrate attention on a particular arc E_{12} with equations (2.1) and inquire what properties it must have if it is to be a minimizing arc. It will be assumed that the functions g, ψ_μ, ϕ_α and the functions $y_i(x)$ defining E_{12} satisfy the hypotheses used by Bliss and Hestenes [VIII, p. 306]. The reader is also referred to their paper for the definition of *admissible set* and *admissible arc.*

Suppose that E_{12} has r corners for values $x = t_\theta$ $(\theta = 1, \cdots, r)$, and $x_1 < t_1 < t_2 < \cdots < t_r < x_2$. To make the notation simpler, let $t_0 = x_1$ and $t_{r+1} = x_2$. At times we shall be interested in the component arcs of E_{12} between corners. We therefore introduce functions $y_i{}^\gamma(x)$ $(\gamma = 1, \cdots, r+1)$ defined by

$$y_i{}^\gamma(x) = y_i(x) \qquad (t_{\gamma-1} \leq x \leq t_\gamma;\ \gamma = 1, \cdots, r+1), \tag{2.5}$$

and denote the sub-arcs of E_{12} thus obtained by $E^\gamma{}_{12}$. A one-parameter family of admissible arcs

$$y_i = y_i(x, b) \qquad x_1(b) \leq x \leq x_2(b)$$

which contains E_{12} for $b = b_0$ and whose end-values satisfy the equations (2.3) may be thought of as a one-parameter family of arcs

$$y_i = y_i{}^\gamma(x, b) \qquad (t_{\gamma-1}(b) \leq x \leq t_\gamma(b);\ \gamma = 1, \cdots, r+1) \tag{2.6}$$

which contains the set $y_i{}^\gamma$ belonging to $E^\gamma{}_{12}$ for $b = b_0$, which satisfies with the set $y'_i{}^\gamma(x, b)$ the equations (2.2) on $t_{\gamma-1}(b) \leq x \leq t_\gamma(b)$, and is such that

$$\begin{aligned} &\text{a)} \quad \psi_\mu[x_1(b), y^1(x_1(b), b), x_2(b), y^{r+1}(x_2(b), b)] = 0 \\ &\text{b)} \quad \chi_i[\theta \mid y] \equiv y_i{}^\theta(t_\theta(b), b) - y_i{}^{\theta+1}(t_\theta(b), b) = 0 \quad (\theta = 1, \cdots, r). \end{aligned} \tag{2.7}$$

The equations of variation of the differential equations (2.2) are

$$\Phi_\alpha[x, \eta, \eta'] \equiv \phi_{\alpha y'_i}\eta'_i + \phi_{\alpha y_i}\eta_i = 0 \qquad (\alpha = 1, \cdots, m), \tag{2.8}$$

where the coefficients $\phi_{\alpha y'_i}$, $\phi_{\alpha y_i}$ have as arguments the functions (2.1) defining E_{12}.

If we set

$$\xi_s = x_{sb}(b_0),\ \tau_\theta = t_{\theta b}(b_0),\ \eta_i = y_{ib}(x, b_0),\ (s = 1, 2;\ \theta = 1, \cdots, r),$$

then the equations of variation corresponding to (2.7) are

$$\begin{aligned} &\text{a)}\quad \Psi_\mu[\xi, \eta] \equiv (\psi_{\mu x_1} + \psi_{\mu y_{i1}}y'_{i1})\xi_1 + (\psi_{\mu x_2} + \psi_{\mu y_{i2}}y'_{i2})\xi_2 \\ &\qquad\qquad + \psi_{\mu y_{i1}}\eta_i(x_1) + \psi_{\mu y_{i2}}\eta_i(x_2) = 0, \\ &\text{b)}\quad \mathbf{X}_i[\theta \mid \tau, \eta] \equiv \{y'_i(t_\theta^-) - y'_i(t_\theta^+)\}\tau_\theta + \eta_i(t_\theta^-) - \eta_i(t_\theta^+) = 0.^* \end{aligned} \tag{2.9}$$

The corresponding variation of $g[x_1, y(x_1), x_2, y(x_2)]$ will be denoted by $G[\xi_1, \eta(x_1), \xi_2, \eta(x_2)]$.

A set of constants ξ_s, τ_θ and functions $\eta_i(x)$ will be called *a set of quasi-admissible variations* if on each interval $t_{\gamma-1} < x < t_\gamma$ the functions $\eta_i(x)$ coincide with functions $\eta_i^\gamma(x)$ which are continuous, consist of a finite number of pieces with continuous derivatives, and satisfy equations (2.8) on $t_{\gamma-1} \leq x \leq t_\gamma (\gamma = 1, \cdots, r+1)$. A set of quasi-admissible variations will be called *a set of generalized admissible variations* if equations (2.9b) are satisfied. It follows readily that a set of generalized admissible variations is a set of admissible variations as defined by Bliss, if and only if $\tau_\theta = 0$ $(\theta = 1, \cdots, r)$.

If $\xi_{s\sigma}, \tau_{\theta\sigma}, \eta_{i\sigma}$ $(\sigma = 1, \cdots, P+1 = p + rn + 1)$ are $P+1$ sets of quasi-admissible variations, it follows by a method similar to that used by Bliss [III, p. 691] that there exists a family

$$\begin{aligned} &y_i = y_i(x, b_1, \cdots, b_{P+1}), \\ &t_{\gamma-1}(b_1, \cdots, b_{P+1}) \leq x \leq t_\gamma(b_1, \cdots, b_{P+1}) \qquad (\gamma = 1, \cdots, r+1) \end{aligned} \tag{2.10}$$

satisfying equations (2.2), containing E_{12} for $b_\sigma = 0$, and having the sets $\xi_{s\sigma}, \tau_{\theta\sigma}, \eta_{i\sigma}$ as its variations along E_{12} with respect to the parameters b_σ. When the equations (2.10) are substituted in the functions g, ψ_μ, χ_i, these become functions of $b_1, \cdots, b_{P+1}$. The following necessary condition may then be proved in the manner used by Bliss to establish the first necessary condition for the problems of Mayer [I, p. 311] and Lagrange [III, p. 683].

I. The First Necessary Condition. *For every minimizing arc* E_{12}

---

* Throughout this paper, the repetition of a subscript in an expression will indicate summation with respect to that subscript over its range of definition. The + and − signs attached to the argument of a function will be used to indicate the right-hand and left-hand limits of the function at the indicated value of the argument.

for the problem of Mayer with corners at $x = t_\theta$ $(\theta = 1, \cdots, r)$ there exists a function $F = \lambda_a(x)\phi_a$ such that between corners of $\dot{E}_{12}$,

$$(2.11) \qquad (d/dx)F_{y'_i} - F_{y_i} = 0, \qquad \phi_a = 0,$$

and such that

$$(2.12) \qquad F_{y'_i}(x_2)\eta_i(x_2) - F_{y'_i}(x_1)\eta_i(x_1) + \lambda_0 G[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] = 0$$

for every set $\xi_1, \eta_i(x_1), \xi_2, \eta(x_2)$ satisfying the equations $\Psi_\mu[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] = 0$. The multipliers $\lambda_a(x)$ coincide on each segment $t_{\gamma-1} < x < t_\gamma$ with functions $\lambda_a^\gamma(x)$ which are continuous on $t_{\gamma-1} \leq x \leq t_\gamma$; furthermore, the functions $F_{y'_i}$, $F - y'_i F_{y'_i}$ are continuous at the corner points $x = t_\theta$ $(\theta = 1, \cdots, r)$.

This first necessary condition differs from that obtained in the usual fashion [IX, p. 480] in that it includes the additional condition that the function $F - y'_i F_{y'_i}$ is continuous at the corners of a minimizing arc.*

If E_{12} is an admissible arc which satisfies between corners the equations (2.11), and which is non-singular, that is, at each element (x, y, y', λ) of E_{12} the determinant

$$R = \begin{vmatrix} F_{y'_i y'_j} & \phi_{a y'_i} \\ \phi_{\beta y'_j} & 0 \end{vmatrix}$$

is different from zero, then the functions $y_i(x)$ and $\lambda_a(x)$ are of classes C'' and C' at least between corners. In this case E_{12} consists of a finite number of extremal arcs and is called a broken extremal. A broken extremal such that the functions $F_{y'_i}$ and $F - y'_i F_{y'_i}$ are continuous at its corners will be called an *extremaloid.*

The admissible arc E_{12} is normal if there exist p sets of admissible variations $\xi_{s\nu}, \eta_{i\nu}$ $(s = 1, 2;\ \nu = 1, \cdots, p)$ such that

$$(2.13) \qquad |\Psi_\mu[\xi_\nu, \eta_\nu]| \neq 0.$$

E_{12} is seen to be normal if and only if there exist P sets of quasi-admissible variations $\xi_{\cdot\pi}, \tau_{\theta\pi}, \eta_{i\pi}$ $(\pi = 1, \cdots, P)$ such that

$$\begin{vmatrix} \Psi_\mu[\xi_\pi, \eta_\pi] \\ X_i[1 \mid \tau_\pi, \eta_\pi] \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ X_i[r \mid \tau_\pi, \eta_\pi] \end{vmatrix} \neq 0.\dagger$$

* The continuity of $F - y'_i F_{y'_i}$ has been derived in another manner by Bliss and Hestenes, who have also discussed the dependence of the continuity of this function upon the continuity of $F_{y'_i}$. For abstract of their paper, see *Bulletin of the American Mathematical Society*, vol. 39 (1933), p. 341.

† One such set is obtained as follows: let $\tau_{\theta\pi} = 0$ $(\pi = 1, \cdots, P;\ \theta = 1, \cdots, r)$,

The admissible arc E_{12} is *normal on the sub-interval* $x'x''$ if there exist $2n-1$ sets of admissible variations ξ_{sk}, $\eta_{ik}(x)$ $(k=1,\cdots,2n-1)$ such that the matrix

$$\left\| \begin{matrix} \eta_{ik}(x') \\ \eta_{ik}(x'') \end{matrix} \right\|$$

has rank $2n-1$.

The method used by Bliss to prove the corresponding result for the problem of Lagrange [III, p. 695], gives the following result.

LEMMA 2.1. *If a set of generalized admissible variations* $\xi_s, \tau_\theta, \eta_i(x)$ *for a normal admissible arc* E_{12} *satisfy the equations* $\Psi_\mu[\xi,\eta]=0$, *then there exists a one-parameter family of admissible arcs*

$$(2.14) \qquad y_i = y_i^\gamma(x,b) \quad t_{\gamma-1}(b) \leq x \leq t_\gamma(b), \qquad (\gamma=1,\cdots,r+1)$$

satisfying the equations (2.7), *containing* E_{12} *for* $b=b_0$, *and having the set* $\xi_s, \tau_\theta, \eta_i(x)$ *as its variations along* E_{12}.

We have also:

LEMMA 2.2. *If* E_{12} *is an extremaloid, then for every set of generalized admissible variations* $\xi_s, \tau_\theta, \eta_i(x)$ *along* E_{12} *the functions* $\eta_i(x)$ *satisfy the relation*

$$(2.15) \qquad F_{y'_i}\eta_i \Big|_{x=x'}^{x=x''} = 0$$

for every set of values x' *and* x'' *on* x_1x_2.

It then follows readily that at each corner $x=t_\theta$ of E_{12} the functions $\eta_i(x)$ belonging to a set of generalized admissible variations ξ_s, τ_θ, $\eta_i(x)$ satisfy the equation

$$(2.16) \qquad F_{y'_i}(t_\theta^-)\eta_i(t_\theta^-) - F_{y'_i}(t_\theta^+)\eta_i(t_\theta^+) = 0.$$

3. Construction of families of extremaloids. Embedding theorems for extremaloids will be obtained by starting with an extremal sub-arc of the given extremaloid and showing how to proceed past a corner. We shall assume that the given extremaloids E_{12} has corners at $x=t_\theta$ $(\theta=1,\cdots,r)$ and is non-singular. Along an extremal sub-arc $E^\gamma{}_{12}$ of E_{12} whose equations are of the form (2.5), the equations

---

and choose ξ_{sv}, η_{iv} $(v=1,\cdots,p)$ as a set of admissible variations satisfying (2.13); finally choose quasi-admissible variations $\eta_{i,p+(\theta-1)n+j}(x)$ as continuous on $x_1 \leq x \leq t_\theta$, satisfying the conditions $\eta_{i,p+(\theta-1)n+j}(t_\theta)=\delta_{ij}$, and as identically zero on $t_\theta < x \leq x_2$ $(\theta=1,\cdots,r)$.

$$F_{y'_i}(x, y, y', \lambda) = z_i, \quad \phi_a(x, y, y') = 0 \tag{3.1}$$

may be solved for the variables y'_i, λ_a in a neighborhood of the values (x, y, z) on $E^{\gamma}{}_{12}$. The solution of (3.1) has the form

$$y'_i = P_i{}^{\gamma}(x, y, z), \lambda_a = \Lambda_a{}^{\gamma}(x, y, z), \tag{3.2}$$

and has continuous partial derivatives of the first two orders since the first members of the equations (3.1) have such derivatives. The system (2.11) is now equivalent to the system

$$dy_i/dx = P_i{}^{\gamma}(x, y, z), \quad dz_i/dx = F_{y_i}[x, y, P^{\gamma}(x, y, z), \Lambda^{\gamma}(x, y, z)]. \tag{3.3}$$

It then follows that through each element (x_0, y_0, z_0) in a neighborhood of the set of values (x, y, z) on $E^{\gamma}{}_{12}$ there passes a unique solution

$$y_i = Y_i{}^{\gamma}(x, x_0, y_0, z_0), \ z_i = Z_i{}^{\gamma}(x, x_0, y_0, z_0) \tag{3.4}$$

of equations (3.3) for which the functions $Y^{\gamma}{}_i, Y^{\gamma}{}_{ix}, Z^{\gamma}{}_i, Z^{\gamma}{}_{ix}$ have continuous partial derivatives of the first two orders since the second members of (3.3) have such derivatives [See III, § 6 and VIII, § 3].

For (x_0, y_0, z_0) in a suitably restricted neighborhood of $E^1{}_{12}$ we then have a family of extremal arcs of the form (3.4) for $\gamma = 1$ embedding $E^1{}_{12}$. Now along this family the corner equations

$$\begin{aligned} &F_{y'_i}[T, Y^1(T), Y^1{}_x(T), \Lambda^1(T)] - F_{y'_i}[T, Y^1(T), p, \lambda] = 0, \\ &F[T, Y^1(T), Y^1{}_x(T), \Lambda^1(T)] - Y^1{}_{ix}(T) F_{y'_i}[T, Y^1(T), Y^1{}_x(T), \Lambda^1(T)] \\ &\qquad - F[T, Y^1(T), p, \lambda] + p_i F_{y'_i}[T, Y^1(T), p, \lambda] = 0, \\ &\phi_a[T, Y^1(T), p] = 0, \end{aligned} \tag{3.5}$$

where for brevity the arguments (x_0, y_0, z_0) are omitted in writing, have initial solutions $T = t_1, \ p_i = y'_i(t_1{}^+), \ \lambda_a = \lambda_a(t_1{}^+), \ (x_0, y_0, z_0)$ equal to an arbitrary element x_{00}, y_{00}, z_{00} on $E^1{}_{12}$. Since E_{12} is non-singular, the functional determinant of the equations (3.5) with respect to the variables p_i, λ_a, T is seen to be equal to a non-zero multiple of the value of Ω_0 at $x = t_1$, where

$$\Omega_0(x) = F_{x'}{}^{+} + y'_i{}^{-} F_{y_i}{}^{+} - F_{x'}{}^{-} - y'_i{}^{+} F_{y_i}{}^{-}, \tag{3.6}$$

and the superscripts + and — denote, respectively, the right-hand and left-hand limits at a given value of x.

Consequently, if we make the additional assumption that $\Omega_0(t_1) \neq 0$, the equations (3.5) have unique solutions $p_i = p_i(x_0, y_0, z_0), \lambda_a = \lambda_a(x_0, y_0, z_0), T = T_1(x_0, y_0, z_0)$ having $x_0, y_0, z_0, p_i, \lambda_a, T$ in suitably restricted neighborhoods of $x_{00}, y_{00}, z_{00}, y'_i(t_1{}^+), \lambda_a(t_1{}^+), t_1$, respectively, and these solutions are of class C' at least. Application of the embedding theorem to the extremal sub-arc $E^2{}_{12}$ now shows that through each element

$$[T_1(x_0, y_0, z_0), \ Y_i{}^1(T_1(x_0, y_0, z_0), x_0, y_0, z_0), \ p_i(x_0, y_0, z_0), \ \lambda_a(x_0, y_0, z_0)]$$

there passes a unique extremal arc in a neighborhood of E^2_{12}, provided (x_0, y_0, z_0) is sufficiently near (x_{00}, y_{00}, z_{00}). If Ω_0 is assumed to be different from zero at each corner of E_{12} we may proceed past these corners successively and obtain a family of extremaloids

$$(3.7) \quad y_i = Y_i(x, x_0, y_0, z_0), \quad z_i = Z_i(x, x_0, y_0, z_0) \qquad (x_1 \leq x \leq x_2)$$

defined for (x, x_0, y_0, z_0) in a neighborhood of the values corresponding to E_{12} and containing E_{12} for $(x_1, x_2, x_0, y_0, z_0) = (x_{10}, x_{20}, x_{00}, y_{00}, z_{00})$. The functions Y_i are continuous in all their arguments, and $Y_i, Y_{ix}, Y_{ixx}, \lambda_a, \lambda_{ax}$ are of class C' in all their arguments except at the corners; furthermore, the corner manifolds

$$(3.8) \quad x = T_\theta(x_0, y_0, z_0), y_i = Y_i[T_\theta(x_0, y_0, z_0), x_0, y_0, z_0] \quad (\theta = 1, \cdots, r)$$

are of class C'.

In the above proof of the existence of the family (3.7) we have assumed that the element (x_{00}, y_{00}, z_{00}) was on E^1_{12}. If this initial element occurred on some other extremal sub-arc of E_{12} it is clear that a family of the form (3.7) would be obtained by extending the family past the corners of the adjoining sub-arcs successively.

Since each curve (3.7) has an initial set at $x = x_{10}$ we lose none of them if we replace x_0 by the fixed value x_{10}. Furthermore, on account of homogeneity relations satisfied by the functions of the set (3.7), the family is seen to depend essentially upon only $2n - 1$ constants [VIII, p. 310]. If these constants are suitably chosen and denoted by $c_1, \cdots, c_{2n-1}$, the equations of the family take the form

$$(3.9) \qquad y_i = y_i(x, c), \quad z_i = z_i(x, c) \qquad (x_1 \leq x \leq x_2)$$

and contains E_{12} for special values $(x_1, x_2, c) = (x_{10}, x_{20}, c_0)$. In a neighborhood of the values (x, c) defining E_{12} the functions

$$y_i, y_{ix}, y_{ixx}, z_i, z_{ix}, \lambda_a = \Lambda_a[x, y(x, c), z(x, c)], \lambda_{ax}$$

are of class C' in all the arguments except along the corner curves of the family, and for the special values (x_{10}, c_0) the determinant

$$(3.10) \qquad \begin{vmatrix} y_{ic_s} & 0 \\ z_{ic_s} & z_i \end{vmatrix}$$

is different from zero. Finally, the equations of the corner manifolds of the family may be written in the form

$$\mathfrak{M}_\theta(c) : x = t_\theta(c), \quad y_i = y_i(t_\theta(c), c) \qquad (\theta = 1, \cdots, r)$$

and the defining functions are of class C' in a neighborhood of $c = c_0$.

4. Further conditions satisfied by a minimizing arc. The Weierstrass necessary condition states that at each element (x, y, y', λ) of a normal minimizing arc E_{12} the inequality

$$(4.1)\quad \mathcal{E}[x, y, y', Y', \lambda] \equiv F(x, y, Y', \lambda) - F(x, y, y', \lambda) - (Y'_i - y'_i)F_{y'_i}(x, y, y', \lambda) \geq 0$$

must be satisfied for every admissible set $(x, y, Y') \neq (x, y, y')$.* The necessary condition of Clebsch states that at each element (x, y, y', λ) of a normal minimizing arc the inequality

$$F_{y'_i y'_j}(x, y, y', \lambda)\Pi_i\Pi_j \geq 0$$

must be satisfied by every set $(\Pi_i) \neq (0)$ which satisfies the equations $\phi_{ay'_i}\Pi_i = 0$. The conditions of Weierstrass and Clebsch will be denoted by II and III, respectively, while II′ and III′ will be used to denote the conditions strengthened to exclude the equality sign.

At a corner $x = t_\theta$ the corner conditions of the first necessary condition imply

$$(4.2)\quad \mathcal{E}[t_\theta, y(t_\theta), y'(t_\theta^-), y'(t_\theta^+), \lambda(t_\theta^-)] = 0 = \mathcal{E}[t_\theta, y(t_\theta), y'(t_\theta^+), y'(t_\theta^-), \lambda(t_\theta^+)].$$

Let $P_i^\theta(x)$ be a set of continuous functions such that $P_i^\theta(t_\theta) = y'_i(t_\theta^+)$ and which satisfies the equations $\phi_a[x, y^\theta(x), P^\theta(x)] = 0$ in a neighborhood of $x = t_\theta$. By the use of the corner conditions we obtain that the function $\mathcal{E}[x, y(x), y'(x), P^\theta(x), \lambda(x)]$ has at $x = t_\theta$ a left-handed derivative equal in value to $\Omega_0(t_\theta)$.† Similarly, if $Q_i^\theta(x)$ is a set of continuous functions satisfying $Q_i^\theta(t_\theta) = y'_i(t_\theta^-)$ and $\phi_a[x, y^{\theta+1}(x), Q^\theta(x)] = 0$ in a neighborhood of $x = t_\theta$, then $\mathcal{E}[x, y(x), y'(x), Q^\theta(x), \lambda(x)]$ has a right-handed derivative at $x = t_\theta$ equal to $-\Omega_0(t)$. We have, therefore, the following results:

If E_{12} is an extremaloid along which $\mathcal{E} \geq 0$, then $\Omega_0 \leq 0$ at the corners of E_{12}.

If E_{12} is a minimizing arc for the problem of Mayer, then $\Omega_0 \leq 0$ at the corners of E_{12}.

We have also the following property:

If E_{12} is an extremal arc and at an element (x, y, y', λ) of E_{12} the $\mathcal{E}$ function is non-negative, but vanishes for an admissible set (x, y, Y')

---

* See Graves, "On the Weierstrass condition for the problem of Bolza in the calculus of variations," *Annals of Mathematics*, vol. 33 (1932), pp. 261-274.

† The corresponding result for the Lagrange problem in parametric form has been established in the above manner by Hefner [VI].

$\neq (x, y, y')$, *then there exist multipliers* $\bar{\lambda}_a$ *such that the set* $(x, y, Y', \bar{\lambda})$ *satisfies the corner conditions with the set* (x, y, y', λ).

The function $\mathcal{E}[x, y, y', p, \lambda]$ is non-negative for all sets p_i such that (x, y, p) is in $\mathfrak{R}$, $\phi_a[x, y, p] = 0$, and $\| \phi_{ay'_j}(x, y, p) \|$ is of rank m. Therefore, since $p_i = Y'_i$ is a set such that

$$\mathcal{E}[x, y, y', Y', \lambda] = 0, \tag{4.3}$$

it follows that there exist constants d_a satisfying the equations

$$(\lambda_a + d_a)\phi_{ay'_i}(x, y, Y') - \lambda_a \phi_{ay'_i}(x, y, y') = 0.$$

If $\bar{\lambda}_a = \lambda_a + d_a$, the set $(x, y, Y', \bar{\lambda})$ satisfies the conditions

$$F_{y'_i}(x, y, Y', \bar{\lambda}) = F_{y'_i}(x, y, y', \lambda),$$

and from (4.3) it then follows that the other corner conditions are also satisfied.

In view of equations (4.2) it is seen that along an extremaloid which is not an extremal arc condition II′ cannot be satisfied. As a matter of notation, however, an arc of the form (2.1) will be said to satisfy condition II′$_0$ if at each element (x, y, y', λ) on the arc we have

$$\mathcal{E}[x, y, y', Y', \lambda] > 0$$

for every admissible set (x, y, Y') which is distinct from the sets $[x, y, y'(x-)]$ and $[x, y, y'(x+)]$.

5. The second variation. Consider a normal non-singular extremaloid E_{12} with end-values satisfying the equations $\psi_\mu = 0$ and which has corners at $x = t_\theta$ $(\theta = 1, \cdots, r)$. Let $\xi_s, \tau_\theta, \eta_i(x)$ be a set of generalized admissible variations satisfying the equations $\Psi_\mu[\xi, \eta] = 0$. By Lemma 2.1 there is a one-parameter family of admissible arcs (2.14) containing E_{12} for $b = b_0$, satisfying $\psi_\mu = 0$, and having the set $\xi_s, \tau_\theta, \eta_i(x)$ as its variations along E_{12}. If the equations of this family are substituted in the functions $g, \psi_\mu, \chi_i, \phi_a$ we obtain *

$$g_2[\xi, \tau, \eta] = \frac{d^2 g}{db^2}\Big]^{b=b_0} = 2Q[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] + \sum_{\theta=1}^{r} 2H[\theta \mid \tau, \eta(t_\theta^-), \eta(t_\theta^+)] + \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')\,dx \tag{5.1}$$

where

$$2\omega = F_{y'_i y'_j}\eta'_i\eta'_j + 2F_{y'_i y_j}\eta'_i\eta_j + F_{y_i y_j}\eta_i\eta_j,$$

$$H[\theta \mid \tau_\theta, \eta(t_\theta^-), \eta(t_\theta^+)] = [\tau_\theta^2\{F_x + F_{y_i}y'_i\} + 2\tau_\theta F_{y_i}\eta_i]\begin{matrix} x = t_\theta^- \\ x = t_\theta^+ \end{matrix},$$

* The reader is referred to [II] for the corresponding calculation of the second variation along an extremal arc.

and Q is a quadratic form in its arguments whose explicit value will not be written.

6. The accessory boundary value problem and an analogue of Jacobi's condition. We shall now consider the problem of minimizing the expression $g_2[\xi, \tau, \eta]$ in the class of generalized admissible variations which satisfy the equations $\Psi_\mu[\xi, \eta] = 0$, and are normed such that

$$\xi_s\xi_s + \tau_\theta\tau_\theta + \int_{x_1}^{x_2} \eta_i(x)\eta_i(x)\,dx = 1. \tag{6.1}$$

The formulation of this problem is simplified by the introduction of new functions $u_k(x)$ $[k = 1, \cdots, n(r+1)]$ by the equations

$$u_{(\gamma-1)n+i}(x) = \eta_i(t_{\gamma-1} + x[t_\gamma - t_{\gamma-1}]) \quad (0 \leq x \leq 1;\ \gamma = 1, \cdots, r+1). \tag{6.2}$$

The limiting values of the functions $\eta_i(x)$ at the points $x = t_\theta$ are then expressible as end-values of the functions $u_k(x)$ at $x = 0$ and $x = 1$. The integrals in (5.1) and (6.1) are seen to become integrals on $0 \leq x \leq 1$ of quadratic forms in the variables $u_k(x)$ and their derivatives, and the set of m equations (2.8) reduces to a set of $m(r+1)$ linear differential equations in the new variables. Finally, let $r+2$ additional functions be defined by the equations

$$\begin{aligned} &u_{rn+n+s}(0) = \xi_s, \quad u_{rn+n+2+\theta}(0) = \tau_\theta && (s = 1, 2;\ \theta = 1, \cdots, r), \\ &u'_{rn+n+s} = 0 = u'_{rn+n+2+\theta} && (0 \leq x \leq 1). \end{aligned}$$

Our accessory minimum problem described above then reduces to the problem of finding in a class of arcs

$$u_\nu = u_\nu(x) \qquad (0 \leq x \leq 1;\ \nu = 1, \cdots, N = rn + r + n + 2), \tag{6.3}$$

which satisfy a set of $mr + m + r + 2$ ordinary linear homogeneous differential equations of the first order, $rn + p$ linear homegeneous equations in the end-values of these functions at $x = 0$ and $x = 1$, and the condition

$$\sum_{v=0}^{r+2} \{u_{rn+n+v}(0)\}^2 + \int_0^1 \sum_{\gamma=1}^{r+1} [\sum_{i=1}^{n} u^2_{(\gamma-1)n+i}(x)] \cdot [t_\gamma - t_{\gamma-1}]\,dx = 1, \tag{6.4}$$

one which minimizes an expression of the form

$$I_2[u] = 2\mathfrak{Q}[u(0), u(1)] + \int_0^1 2w(x, u, u')\,dx. \tag{6.5}$$

In this last expression $\mathfrak{Q}$ and w are quadratic forms in $u_\nu(0), u_\nu(1)$ and u_ν, u'_ν respectively.

The boundary value problem consisting of the Euler-Lagrange equations and transversality conditions for this reduced form of the accessory minimizing problem is of the form recently considered by the author [VII, and

X, § 6]. If E_{12} is a normal extremaloid which satisfies condition III′, then the conditions of Theorem 6.1 of [X] are satisfied and there exist infinitely many real characteristic numbers $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots$ of the boundary value problem. Furthermore, each characteristic number is determined by a corresponding minimizing property. In particular, λ_1 is the minimum of $I_2[u]$ in the class of arcs (6.3) which satisfy the differential equations, end-conditions, and condition (6.4).

It is more convenient for our use, however, to write the differential equations and boundary conditions of this accessory boundary value problem in terms of the original set of variations $\xi_s, \tau_\theta, \eta_i(x)$ defined on x_1x_2. It is also advantageous to give the differential equations in canonical form. As is customary, let

$$\Omega(x, \eta, \eta', \mu) = \omega(x, \eta, \eta') + \mu_a \Phi_a(x, \eta, \eta').$$

Along a non-singular extremaloid E_{12} the equations

$$\zeta_i = \Omega_{\eta' i}(x, \eta, \eta', \mu), \qquad \Phi_a(x, \eta, \eta') = 0 \tag{6.7}$$

have unique solutions

$$\eta' i = \mathrm{H}_i[x, \eta, \zeta], \qquad \mu_a = \mathrm{M}_a[x, \eta, \zeta] \tag{6.8}$$

which are linear in the variables η_i, ζ_i. It then follows that the characteristic numbers of the accessory boundary value problem are the values of λ corresponding to which there exist sets $\xi_s, \tau_\theta, \eta_i(x), \zeta_i(x)$ which satisfy the following conditions:

(a) the set $\xi_s, \tau_\theta, \eta_i(x)$ is a set of generalized admissible variations;

(b) the functions $\eta_i(x)$, $\zeta_i(x)$ coincide on each segment $t_{\gamma-1} < x < t_\gamma$ with functions which are of class C' on $t_{\gamma-1} \leq x \leq t_\gamma$;

(c) the functions η_i, ζ_i are not all identically zero on x_1x_2, and satisfy on each segment $t_{\gamma-1} < x < t_\gamma$ the equations

$$\eta' i = \mathrm{H}_i[x, \eta, \zeta], \ \zeta' i = \Omega_{\eta_i}(x, \eta, \mathrm{H}[x, \eta, \zeta], \mathrm{M}[x, \eta, \zeta]) - \lambda\eta_i = 0, \ \Phi_a = 0; \tag{6.9}$$

(d) there exist constants

$$d_\mu, d_{\theta|i} \qquad (\mu = 1, \cdots, p;\ \theta = 1, \cdots, r;\ i = 1, \cdots, n)$$

which satisfy the equations

$$\begin{aligned}
&\Psi_\mu[\xi, \eta] = 0 && (\mu = 1, \cdots, p),\\
&\mathrm{X}_i[\theta \mid \tau, \eta] = 0 && (i = 1, \cdots, n;\ \theta = 1, \cdots, r),\\
&Q_{\eta_{i_s}} + d_\mu \Psi_{\mu\eta_{i_s}} + (-1)^s \zeta_i(x_s) = 0,\\
&Q_{\xi_s} + d_\mu \Psi_{\mu\xi_s} - \lambda\xi_s = 0 && (s = 1, 2),\\
&\tau_\theta F_{y_i}(t_\theta^-) + d_{\theta|i} + \zeta_i(t_\theta^-) = 0,\\
&\tau_\theta F_{y_i}(t_\theta^+) + d_{\theta|i} + \zeta_i(t_\theta^+) = 0,\\
&H_{\tau_\theta}[\theta \mid \tau, \eta] + \{y'_i(t_\theta^-) - y'_i(t_\theta^+)\}\, d_{\theta|i} - \lambda\tau_\theta = 0.
\end{aligned} \tag{6.10}$$

It is to be remarked that if $\xi_s, \tau_\theta, \eta_i, \zeta_i$ is a solution of the system (6.9), (6.10), then the multipliers $\mu_a(x)$ defined by (6.8), together with the functions $\eta'_i(x)$, coincide on $t_{\gamma-1} < x < t_\gamma$ with functions which are of class C' on $i_{\gamma-1} \leq x \leq t_\gamma$ $(\gamma = 1, \cdots, r+1)$.

We have the following important theorem.

THEOREM 6.1. *Let E_{12} be a normal admissible arc whose end-values satisfy the equations $\Psi_\mu = 0$, and which satisfies conditions I and III'. Then a necessary and sufficient condition that $g_2[\xi, \tau, \eta] \geq 0$ for all sets of generalized admissible variations $\xi_s, \tau_\theta, \eta_i(x)$ which satisfy the equations $\Psi_\mu = 0$ is that the least characteristic number λ_1 of the corresponding boundary value problem (6.9), (6.10) be non-negative.*

As a corollary to this theorem we have the following result:

IV. THE CONDITION OF MAYER. *For a normal non-singular minimizing arc E_{12} the least characteristic number λ_1 of the corresponding accessory boundary value problem (6.9), (6.10) must satisfy the condition $\lambda_1 \geq 0$.*

As is customary, IV' will be used to denote condition IV with the equality sign excluded.

7. Secondary extremaloids. The differential equations (6.9) reduce for $\lambda = 0$ to the set

$$(7.1)\quad \eta'_i = \mathrm{H}_i[x, \eta, \zeta],\ \zeta'_i = \Omega_{\eta_i}(x, \eta, \mathrm{H}[x, \eta, \zeta], \mathbf{M}[x, \eta, \zeta]),\ \Phi_a(x, \eta, \eta') = 0.$$

This set is called the *accessory system of differential equations*. We shall understand by a *secondary extremaloid* a set of functions $\eta_i(x)$, $\zeta_i(x)$ and constants τ_θ satisfying condition (b) of § 6, which is such that equations (7.1) are satisfied on each segment $t_{\gamma-1} < x < t_\gamma$, and which satisfies with associated constants $d_{\theta|i}$ the corner equations

$$(7.2)\quad \begin{aligned} \mathrm{X}_i[\theta \mid \tau, \eta] &= 0, \\ \mathrm{X}_{n+i}[\theta \mid \tau, \eta, d] &\equiv \tau_\theta F_{y_i}(t_\theta^-) + d_{\theta|i} + \zeta_i(t_\theta^-) = 0, \\ \mathrm{X}_{2n+i}[\theta \mid \tau, \eta, d] &\equiv \tau_\theta F_{y_i}(t_\theta^+) + d_{\theta|i} + \zeta_i(t_\theta^+) = 0, \\ \mathrm{X}_{3n+1}[\theta \mid \tau, \eta, d] &\equiv H_{\tau_\theta}[\theta \mid \tau, \eta] + \{y'_i(t_\theta^-) - y'_i(t_\theta^+)\} d_{\theta|i} = 0 \\ &\qquad (\theta = 1, \cdots, r). \end{aligned}$$

Equations (7.2) are seen to be those of the set (6.10) which apply at the points $x = t_\theta$ with $\lambda = 0$.

The following lemma is of significance:

LEMMA 7.1. *If E_{12} is a normal non-singular extremaloid having $\Omega_0 \neq 0$ at its corners, then corresponding to a secondary extremal arc η_i, ζ_i defined on one of the intervals $t_{\gamma-1} \leq x \leq t_\gamma$ there is a unique secondary extremaloid on $x_1 \leq x \leq x_2$ whose defining functions coincide with η_i, ζ_i on their interval of definition.*

In the $3n+1$ equations of (7.2) corresponding to a given value of θ the determinant of the coefficients of the quantities $\tau_\theta, d_{\theta|i}, \eta_i(t_\theta^-), \zeta_i(t_\theta^-)$, or of the quantities $\tau_\theta, d_{\theta|i}, \eta_i(t_\theta^+), \zeta_i(t_\theta^+)$, is seen to be a non-zero multiple of $\Omega_0(t_\theta)$. Lemma 7.1 then follows readily from the hypothesis $\Omega_0(t_\theta) \neq 0$ and the existence theorem for the system (7.1).

In particular, it is to be noted that $\eta_i \equiv 0$, $\zeta_i \equiv z_i(x)$ is a secondary extremaloid for which $\tau_\theta = 0$ $(\theta = 1, \cdots, r)$.

8. The determination of conjugate points. As before, we consider a normal non-singular extremaloid E_{12} with corners at $x = t_\theta$ $(\theta = 1, \cdots, r)$.

DEFINITION OF CONJUGATE POINT. *A value x_3 is said to define a point 3 conjugate to the point 1 on E_{12} if there exists a secondary extremaloid $\tau_\theta = \sigma_\theta, \eta_i = u_i(x), \zeta_i = v_i(x)$ whose functions $u_i(x)$ are not all identically zero on $x_1 x_3$, and satisfy with constants c, d the conditions*

$$u_i(x_1) = 0 = cu_i(x_3^-) + du_i(x_3^+) \qquad (cd \geq 0,\ c + d \neq 0). \tag{8.1}$$

If x_3 does not define a corner of E_{12} the conditions (8.1) become $u_i(x_1) = 0 = u_i(x_3)$.

The following lemma is a consequence of equations (7.2) and (3.6), and the proof will be omitted. The results will be used in establishing certain properties of conjugate points.

LEMMA 8.1. *Let $\tau_0, \eta_i(x), \zeta_i(x)$ be a secondary extremaloid which at a particular corner point $x = t_\theta$ of E_{12} satisfies with constants c and d the relation*

$$c\eta_i(t_\theta^-) + d\eta_i(t_\theta^+) = 0 \qquad (c + d \neq 0). \tag{8.2}$$

Then

$$2H[\theta \mid d\tau_\theta/(c+d), \eta(t_\theta^-), 0] + \eta_i(t_\theta^-)\zeta_i(t_\theta^-) = cd\tau_\theta^2\Omega_0(t_\theta)/(c+d)^2, \tag{8.3}$$

$$2H[\theta \mid c\tau_\theta/(c+d), 0, \eta(t_\theta^+)] - \eta_i(t_\theta^+)\zeta_i(t_\theta^+) = cd\tau_\theta^2\Omega_0(t_\theta)/(c+d)^2. \tag{8.4}$$

A generalization of the result obtained by Bliss and Hestenes [VIII, p. 315] for extremal arcs in the special Mayer problem for which $p = 2n+1$ is given by the following necessary condition.

IV$_0$. *Let E_{12} be a normal non-singular extremaloid, normal on every sub-interval $x_3 x_2$ of E_{12}, and having $\Omega_0 \neq 0$ at each of its corners. If E_{12} is a minimizing arc for the problem of Mayer here proposed, then between 1 and 2 on E_{12} there can be no points 3 conjugate to 1.*

If there were a secondary extremaloid σ_θ, u_i, v_i whose functions $u_i(x)$ satisfied conditions (8.1) but were not all identically zero on $x_1 x_3$, then let

$$\eta_i(x) = u_i(x) \quad (x_1 \leq x \leq x_3), \quad \eta_i(x) \equiv 0 \qquad (x_3 < x \leq x_3). \tag{8.5}$$

In case either $u_i(x_3^-) = 0$ or $u_i(x_3^+) = 0$ $(i = 1, \cdots, n)$, corresponding constants τ_θ would be defined as follows: if $u_i(x_3^-) = 0$ $(i = 1, \cdots, n)$ then $\tau_\theta = \sigma_\theta$ if $t_\theta < x_3$ and $\tau_\theta = 0$ if $t_\theta \geq x_3$; if the values $u_i(x_3^-)$ were not all zero but $u_i(x_3^+) = 0$ $(i = 1, \cdots, n)$, then $\tau_\theta = \sigma_\theta$ if $t_\theta \leq x_3$ and $\tau_\theta = 0$ if $t_\theta > x_3$. These values, together with the functions η_i of (8.5) and the constants $\xi_s = 0$, would be such that $g_2[\xi, \tau, \eta] = 0$, and would be therefore a minimizing set for the accessory minimum problem of § 6. There would then exist multipliers $\mu_a(x)$ and functions $\zeta_i(x)$ defined by (6.7) which were of class C' and satisfied with the functions (8.5) equations (7.1) on every sub-interval of x_1x_2 which did not contain any of the points x_3, t_θ $(\theta = 1, \cdots, r)$. Furthermore, the functions $\zeta_i(x)$ would be continuous between corners of E_{12} and equations (7.2) would be satisfied by suitable constants $d_{\theta|i}$. In view of the normality of E_{12} on x_3x_2, there would then exist a constant k such that $\zeta_i(x) = kz_i(x)$ on x_3x_2, and as a consequence of the continuity of $\zeta_i(x)$ between corners of E_{12} and Lemma 7.1 it would follow that $u_i(x) \equiv 0$ on x_1x_3, which is a contradiction.

There still remains the possibility of x_3 defining a corner point t_{θ^*} of E_{12} and the functions of neither of the sets $u_i(x_3^-)$, $u_i(x_3^+)$ being all zero. In this case it would follow from the equations $X_i[\theta^* \mid \sigma, u] = 0$ that $\sigma_{\theta^*} \neq 0$. With the functions $\eta_i(x)$ of (8.5) we would then associate constants ξ_s, τ_θ as follows:

$$\xi_s = 0 \ (s = 1, 2); \quad \tau_\theta = \sigma_\theta \text{ if } \theta < \theta^*, \quad \tau_\theta = 0 \text{ if } \theta > \theta^*, \quad \tau_{\theta^*} = d\sigma_{\theta^*}/(c + d).$$

It would then follow from Lemma 8.1 that

$$g_2[\xi, \tau, \eta] = cd\sigma^2_{\theta^*}\Omega_0(t_{\theta^*})/(c + d)^2 < 0.$$

Since this is impossible if E_{12} is a minimizing arc, the necessary condition IV_0 is therefore established.

The notation IV'_0 will be used to denote the condition that there exist no value x on $x_1 < x \leq x_2$ which defines a point conjugate to the point 1 on E_{12}. It follows readily that condition IV' implies IV'_0.

If $\bar\eta_i, \bar\zeta_i$ and $\bar u_i, \bar v_i$ are a pair of secondary extremal arcs on a segment $t_{\gamma-1} < x < t_\gamma$, then the expression $\bar\eta_i\bar v_i - \bar\zeta_i\bar u_i$ is equal to a constant on this segment [III, p. 738]; in particular, if this constant is zero the secondary extremal arcs are said to be conjugate. Let $\tau_\theta, \eta_i, \zeta_i$ and σ_θ, u_i, v_i denote the secondary extremaloids on x_1x_2 whose functions η_i, ζ_i and u_i, v_i coincide with $\bar\eta_i, \bar\zeta_i$ and $\bar u_i, \bar v_i$, respectively, on their interval of definition. From the conditions (7.2) satisfied by these secondary extremaloids it follows that the function $\eta_iv_i - \zeta_iu_i$ is constant on x_1x_2. Hence, if on any segment $t_{\gamma-1} < x < t_\gamma$ the component secondary extremal arcs of two secondary extremaloids are conjugate, then on each segment $t_{\gamma-1} < x < t_\gamma$ $(\gamma = 1, \cdots, r + 1)$

the component secondary extremal arcs are conjugate; two such secondary extremaloids are said to be conjugate.

The following lemma will be of use in the proof of the succeeding theorem.

LEMMA 8.2. *Let* $\sigma_{\theta j}, u_{ij}(x)$ $(j=1,\cdots,n)$ *be arbitrary generalized admissible variations on* x_1x_2. *Then for arbitrary values of* c *and* d,

$$|cu_{ij}(x_3^-)+du_{ij}(x_3^+)| = (c+d)^{n-1}[c\,|u_{ij}(x_3^-)|+d\,|u_{ij}(x_3^+)|] \qquad (x_1\leq x\leq x_2). \tag{8.6}$$

If x_3 does not define a corner of E_{12}, then equation (8.6) is trivial. If $x_3=t_\theta$ $(\theta=1,\cdots,r)$ the result may be proved in an elementary manner in view of the equations $X_i[\theta\,|\,\sigma,u]=0$.

The following theorem is fundamental.

THEOREM 8.1. *Let* E_{12} *be a normal non-singular extremaloid which has* $\Omega_0\neq 0$ *at its corners, and which is normal on every sub-interval* $x_1x_3(x_1<x_3\leq x_2)$. *If* $\tau_{\theta j},\eta_{ij}(x),\zeta_{ij}(x)$ $(j=1,\cdots,n)$ *are secondary extremaloids determined by the initial conditions*

$$\begin{aligned} &\eta_{ih}(x_1)=0, \quad \eta_{in}(x_1)=z_i(x_1) && (h=1,\cdots,n-1),\\ &|\zeta_{ih}(x_1)z_i(x_1)|\neq 0, \quad \zeta_{in}(x_1)=0, \end{aligned} \tag{8.7}$$

then a value $x_3\neq x_1$ *defines a point* 3 *conjugate to* 1 *on* E_{12} *if and only if* $|\eta_{ij}(x_3^-)|\cdot|\eta_{ij}(x_3^+)|\leq 0$.

On the assumption that E_{12} is normal on every sub-interval $x_1x_3(x_1<x_3\leq x_2)$, it is seen that an arc $\eta_i=\eta_{ij}(x)b_j$ is identically zero on x_1x_3 if and only if $b_j=0$ $(j=1,\cdots,n)$. It is also to be remarked that if for such a point x_3 there exist constants c, d, b_j such that

$$[c\eta_{ij}(x_3^-)+d\eta_{ij}(x_3^+)]b_j=0 \qquad (c+d\neq 0), \tag{8.8}$$

then $b_n=0$. This result is immediate, since by Lemma 2.2 we have

$$\begin{aligned} 0=z_i(x_3)[c\eta_{ij}(x_3^-)+d\eta_{ij}(x_3^+)]b_j &= (c+d)z_i(x_1)\eta_{ij}(x_1)b_j\\ &= (c+d)z_i(x_1)z_i(x_1)b_n. \end{aligned}$$

Now if x_3 is a point such that $|\eta_{ij}(x_3^-)|\cdot|\eta_{ij}(x_3^+)|\leq 0$, there exist constants c, d satisfying the relation

$$c\,|\eta_{ij}(x_3^-)|+d\,|\eta_{ij}(x_3^+)|=0 \qquad (cd\geq 0,\ c+d\neq 0). \tag{8.9}$$

As a consequence of Lemma 8.2 there then exist constants b_j not all zero, but with b_n necessarily zero, which satisfy equations (8.8) with the values c, d determined by (8.9). We have therefore established that if the de-

terminant $|\eta_{ij}(x)|$ vanishes or changes sign at $x = x_3$, then x_3 defines a point 3 conjugate to 1 on E_{12}.

Conversely, it is seen that every secondary extremaloid $\tau_\theta, \eta_i, \zeta_i$ which satisfies the conditions $\eta_i(x_1) = 0$ is of the form

$$\eta_i = \eta_{ih}(x)b_h, \quad \zeta_i = \zeta_{ih}(x)b_h + z_i(x)b_n, \quad \tau_\theta = \tau_{\theta h}b_h.$$

If x_3 defines a point 3 conjugate to 1 on E_{12} there exist constants c, d such that $|c\eta_{ij}(x_3^-) + d\eta_{ij}(x_3^+)| = 0$, $cd \geq 0$, $c + d \neq 0$, and as a consequence of Lemma 8.2 we have $|\eta_{ij}(x_3^-)| \cdot |\eta_{ij}(x_3^+)| \leq 0$.

9. An auxiliary theorem. In this section we shall prove the following auxiliary theorem which is fundamental in the construction of a family of extremaloids for the problem of Mayer.

THEOREM 9.1. *Let E_{12} be a normal extremaloid, normal on every subinterval x_1x_3, which satisfies conditions III′, IV′$_0$ and for which $\Omega_0 < 0$ at the corner values $x = t_\theta$ $(\theta = 1, \cdots, r)$. Then there exists a set of n secondary extremaloids $S_{\theta j}$, $U_{ij}(x)$, $V_{ij}(x)$ which are mutually conjugate in pairs and such that the determinant $|U_{ij}(x)|$ does not vanish or change sign on the interval $x_1 \leq x \leq x_2$.*

Preliminary to the proof of the above theorem we shall establish two lemmas. Consider a system of secondary extremaloids $\tau_{\theta j}$, $\eta_{ij}(x)$, $\zeta_{ij}(x)$ defined by initial conditions (8.7), and such that $\zeta_{ih}(x_1)z_i(x_1) = 0$ $(h = 1, \cdots, n-1)$. The members of this family are then seen to be mutually conjugate in pairs. Furthermore, since E_{12} satisfies the condition IV′$_0$ it follows that the determinant $|\eta_{ij}(x)|$ does neither vanish nor change sign on $x_1 < x \leq x_2$. If the matrices $\|\tau_{\theta j}\|$, $\|\eta_{ij}(x)\|$, $\|\zeta_{ij}(x)\|$ are multiplied on the right by the inverse of the matrix $\|\eta_{ij}(x_2)\|$ a new conjugate system $\sigma_{\theta j}$, $u_{ij}(x)$, $v_{ij}(x)$ is obtained which satisfies the conditions $u_{ij}(x_2) = \delta_{ij}$, $v_{ij}(x_2) = B_{ij} = v_{ji}(x_2)$ $(i, j = 1, \cdots, n)$. This conjugate system has the following properties:

LEMMA 9.1. *Let ξ_s, τ_θ, $\eta_i(x)$ be an arbitrary set of generalized admissible variations, and define on $x_1 < x \leq x_2$ functions $a_j(x)$ by the relation $\eta_i(x) = u_{ij}(x)a_j(x)$, where the functions $u_{ij}(x)$ belong to the conjugate set of secondary extremaloids $\sigma_{\theta j}$, $u_{ij}(x)$, $v_{ij}(x)$ defined above. Then*

$$(9.1) \qquad [\tau_\theta - \sigma_{\theta j}a_j(t_\theta^-)][\tau_\theta - \sigma_{\theta j}a_j(t_\theta^+)] \geq 0 \qquad (\theta = 1, \cdots, r),$$

$$(9.2) \qquad 2H[\theta \mid \tau, \eta] - [\eta_i(x)v_{ij}(x)a_j(x)]_{t_\theta^-}^{t_\theta^+} + [\tau_\theta - \sigma_{\theta j}a_j(t_\theta^-)][\tau_\theta - \sigma_{\theta j}a_j(t_\theta^+)]\Omega_0(t_\theta) = 0.$$

The equations $X_i[\theta \mid \tau, \eta] = 0$, $X_i[\theta \mid \sigma_j, u_j] = 0$ are seen to imply that if for a given value of θ the functions $a_j(t_\theta^-) - a_j(t_\theta^+)$ are all zero then the quantity $\tau_\theta - \sigma_{\theta j} a_j(t_\theta^-)$ is also zero. It also follows that the functions $u_i(x \mid \theta) = u_{ij}(x)[a_j(t_\theta^-) - a_j(t_\theta^+)]$ satisfy the relations

$$[\tau_\theta - \sigma_{\theta j} a_j(t_\theta^-)] u_i(t_\theta^- \mid \theta) - [\tau_\theta - \sigma_{\theta j} a_j(t_\theta^+)] u_j(t_\theta^+ \mid \theta) = 0.$$

Relation (9.1) is then seen to follow as a consequence of the condition that E_{12} satisfies IV'_0.

To prove (9.2) it is only necessary to verify that the expression on the left-hand side of the equation is identical with the rather complicated quantity

$$\begin{aligned}
&\{\tau_\theta[F_{y_i}(t_\theta^-) + F_{y_i}(t_\theta^+)] - F_{y_i}(t_\theta^+)\sigma_{\theta k} a_k(t_\theta^+) - F_{y_i}(t_\theta^-)\sigma_{\theta k} a_k(t_\theta^-)\} X_i[\theta \mid \tau, \eta] \\
&\quad + \{a_j(t_\theta^-) + a_j(t_\theta^+)\}\{-X_i[\theta \mid \tau, \eta] d_\theta|_{i,j} - X_{2n+i}[\theta \mid \sigma_j, u_j, d_j]\eta_i(t_\theta^+) \\
&\quad + X_{3n+1}[\theta \mid \sigma_j, u_j, d_j]\tau_\theta\} + a_j(t_\theta^-)\{X_{n+i}[\theta \mid \sigma_j, u_j, d_j]\eta_i(t_\theta^-) \\
&\quad - F_{y_i}(t_\theta^+) X_i[\theta \mid \sigma_j, u_j]\tau_\theta\} - \tau_\theta F_{y_i}(t_\theta^-) X_i[\theta \mid \sigma_j, u_j] a_j(t_\theta^+) \\
&\quad + a_k(t_\theta^+)\{[d_{\theta|i,k} + \sigma_{\theta k} F_{y_i}(t_\theta^+)] X_i[\theta \mid \sigma_j, u_j] \\
&\quad + F_{y_i}(t_\theta^-) X_i[\theta \mid \sigma_k, u_k]\sigma_{\theta j} \\
&\quad + u_{ij}(t_\theta^+) X_{2n+i}[\theta \mid \sigma_k, u_k, d_k] \\
&\quad - X_{3n+1}[\theta \mid \sigma_k, u_k, d_k]\sigma_{\theta j}\} a_j(t_\theta^-).
\end{aligned}$$

LEMMA 9.2. *Suppose σ_θ, $u_i(x)$, $v_i(x)$ is a secondary extremaloid and there is a point x_3 on $x_1 \leq x_3 < x_2$ such that*

$$cu_i(x_3^-) + du_i(x_3^+) = 0 \qquad cd \geq 0,\ c + d \neq 0. \tag{9.3}$$

Now define an arc η_i by the equations

$$\eta_i = 0 \quad (x_1 \leq x < x_3), \qquad \eta_i = u_i(x) \quad (x_3 < x \leq x_2)$$

and determine associated constants ξ_s, τ_θ as follows: ξ_s $(s = 1, 2)$ arbitrary; $\tau_\theta = 0$ if $t_\theta < x_3$, $\tau_\theta = \sigma_\theta$ if $t_\theta > x_3$; if x_3 coincides with a corner point t_{θ^*} of E_{12} then $\tau_{\theta^*} = c\sigma_{\theta^*}/(c + d)$. Then the set of generalized admissible variations ξ_s, τ_θ, $\eta_i(x)$ so defined satisfies the inequality

$$g_2[\xi, \tau, \eta] - \eta_i(x_2) B_{ij} \eta_j(x_2) - 2Q[\xi, \eta] \geq 0. \tag{9.4}$$

Consider first the case when $x_1 < x_3 < x_2$. It then follows by the Clebsch transformation of the second variation [III, p. 738] that

$$\begin{aligned}
g_2[\xi, \tau, \eta] = \eta_i(x_2) B_{ij} \eta_j(x_2) + \sum_{\theta=1}^{r} 2H[\theta \mid \tau, \eta] - \eta_i(x) v_{ij}(x) a_j(x) \Big|_{t_\theta^-}^{t_\theta^+} \\
+ 2Q[\xi, \eta] + \int_{x_1}^{x_2} F_{y'_i y'_j}(\eta'_i - u'_{ik} a_k)(\eta'_j - u'_{jl} a_l)\, dx.
\end{aligned}$$

The expression (9.4) is then seen to be non-negative in view of (9.1), (9.2), condition III′, and the hypothesis that $\Omega_0 < 0$ at the corners of E_{12}. Suppose

now that $x_3 = x_1$. It is to be remarked first that we may assume without loss of generality that $\zeta_i(x_1)z_i(x_1) = 0$, since this additional property may be obtained by adding a suitable multiple of the set $\xi_s = 0$, $\tau_\theta = 0$, $\eta_i \equiv 0$, $\zeta_i(x) = z_i(x)$, and such modification does not change the value of the expression in (9.4). The modified secondary extremaloid, however, is seen to be expressible linearly in terms of the members of the family $\sigma_{\theta j}$, u_{ij}, v_{ij}, and by direct integration it is found that $g_2[\xi, \tau, \eta] - 2Q[\xi, \eta] = \eta_i(x_2)B_{ij}\eta_j(x_2)$. Hence the lemma is established.

In order to establish Theorem 9.1 it will now be proved that the set of mutually conjugate secondary extremaloids $S_{\theta j}$, $U_{ij}(x)$, $V_{ij}(x)$ having the initial values

$$U_{ij}(x_2) = \delta_{ij}, \qquad V_{ij}(x_2) = D_{ij} = B_{ij} - \delta_{ij} \tag{9.5}$$

has the property that the determinant $|U_{ij}(x)|$ neither vanishes nor changes sign on $x_1 \leq x \leq x_2$. In the first place, $|U_{ij}(x_2)| = 1$. If now $|U_{ij}(x)|$ were to vanish or change sign for a value $x_3 (x_1 \leq x_3 < x_2)$, there would exist constants a_j not all zero such that the secondary extremaloid $\sigma_\theta = S_{\theta j}a_j$, $u_i = U_{ij}a_j$, $v_i = V_{ij}a_j$ satisfied at x_3 the condition (9.3). Then consider the set ξ_s, τ_θ, η_i, ζ_i of Lemma 9.2. In case x_3 did not correspond to a corner of E_{12}, it could be proved by direct integration that

$$g_2[\xi, \tau, \eta] - \eta_i(x_2)B_{ij}\eta_j(x_2) - 2Q[\xi, \eta] = a_iD_{ij}a_j - a_iB_{ij}a_j = -a_ia_i < 0.$$

In case x_3 coincided with a corner value $x = t_{\theta\bullet}$ of E_{12}, then by direct integration and the use of relation (8.4) it would follow that

$$\begin{aligned} g_2[\xi, \tau, \eta] &- \eta_i(x_2)B_{ij}\eta_j(x_2) - 2Q[\xi, \eta] \\ &= a_iD_{ij}a_j - a_iB_{ij}a_j + cd\tau_{\theta\bullet}{}^2\Omega_0(t_{\theta\bullet})/(c+d)^2, \leq -a_ia_i < 0, \end{aligned}$$

in view of the assumption that $\Omega_0 < 0$ at the corners of E_{12}. In either case we would have a contradiction to the result of Lemma 9.2. Hence the determinant $|U_{ij}(x)|$ neither vanishes nor changes sign on $x_1 \leq x \leq x_2$ and Theorem 9.1 is established.

10. Definition of a field and a fundamental sufficiency theorem. A field with r discontinuities is defined to be a region $\mathfrak{F}$ in xy-space containing only interior points, having associated with it r corner manifolds

$$\mathfrak{M}_\theta : x = T_\theta(a_1, \cdots, a_n),\ y_i = Y_{\theta|i}[T_\theta(a_1, \cdots, a_n), a_1, \cdots, a_n], \quad (\theta = 1, \cdots, r) \tag{10.1}$$

and functions $p_i(x, y)$, $l_\alpha(x, y)$ $(i = 1, \cdots, n;\ \alpha = 1, \cdots, m)$ with the following properties:

(a) the corner manifolds $\mathfrak{M}_\theta$ are non-singular, have no multiple points, do not intersect each other and each $\mathfrak{M}_\theta$ divides $\mathfrak{F}$ into two parts;

(b) the functions $p_i(x, y)$, $l_a(x, y)$ are continuous and have continuous first derivatives between corner manifolds $\mathfrak{M}_\theta$ in $\mathfrak{F}$ and approach finite limits on each side of each $\mathfrak{M}_\theta$;

(c) the two limits p_i^- and p_i^+ of the functions p_i at a point of a corner manifold $\mathfrak{M}_\theta$ are such that the sets p_i^- and p_i^+ always determine directions on the same side of $\mathfrak{M}_\theta$ and never tangent to it;

(d) the sets $[x, y, p(x, y)]$ for (x, y) in $\mathfrak{F}$ are all admissible;

(e) the functions

$$F_{y'_i}[x, y, p(x, y), l(x, y)],$$
$$F[x, y, p(x, y), l(x, y)] - p_i(x, y) F_{y'_i}[x, y, p(x, y), l(x, y)]$$

are continuous in $\mathfrak{F}$;

(f) the Hilbert integral

$$I^* = \int \{F[x, y, p, l]dx + (dy_i - p_i dx) F_{y'_i}[x, y, p, l]\}$$

formed with these functions is independent of the path in $\mathfrak{F}$. At a point (x, y) of a corner manifold of $\mathfrak{F}$ it is to be understood that the set of values (x, y, p, l) appearing in the integrand of I^* are replaced by either the set (x, y, p^-, l^-) or the set (x, y, p^+, l^+). From (e) it follows that the value of the integrand is the same for either set of limiting values.

It may be proved in the usual manner that in the field between corner manifolds $\mathfrak{M}_\theta$ the solutions of the differential equations

$$dy_i/dx = p_i(x, y)$$

are extremal arcs. Condition (e) above then implies that these may be pieced together to form extremaloids, which are called the extremaloids of the field. The value of I^* is seen to be zero along every extremaloid of the field.

Sufficient conditions for the problem of Mayer here discussed are obtained by the same method that Hestenes [IX] has used for the case of extremal arcs, and makes use of the following auxiliary Mayer problem. On the assumption that E_{12} is a normal extremaloid which satisfies the necessary condition I it is seen that the matrix

$$\left\| \begin{matrix} g_{x_1} & g_{y_{i1}} & g_{x_2} & g_{y_{i2}} \\ \psi_{\mu x_1} & \psi_{\mu y_{i1}} & \psi_{\mu x_2} & \psi_{\mu y_{i2}} \end{matrix} \right\| \tag{10.2}$$

has rank $p+1$. Let $\psi_v[x_1, y_1, x_2, y_2]$ $(v = p+1, \cdots, 2n+1)$ be $2n+1-p$ functions possessing continuous first and second partial derivatives in a neighborhood of the end-values $[x_{10}, y_{10}, x_{20}, y_{20}]$ belonging to E_{12}, vanishing at these values, and for these values having the determinant

$$(10.3) \qquad \begin{vmatrix} g_{x_1} & g_{y_{i1}} & g_{x_2} & g_{y_{i2}} \\ \psi_{\rho x_1} & \psi_{\rho y_{i1}} & \psi_{\rho x_2} & \psi_{\rho y_{i2}} \end{vmatrix}$$

different from zero. The new set of end conditions $\psi_\rho = 0$ $(\rho = 1, \cdots, 2n+1)$ defines an auxiliary problem of the type discussed by Bliss and Hestenes [VIII], and a minimizing extremaloid E_{12} for the original Mayer problem is also a minimizing extremaloid for this auxiliary problem. Furthermore, if E_{12} is not only normal, but also normal on x_1x_2, then it is normal with respect to the conditions $\psi_\rho = 0$ $(\rho = 1, \cdots, 2n+1)$ as defined above [see IX, p. 484].

For this auxiliary problem of Mayer we may prove by the same methods that Bliss and Hestenes have used for extremal arcs [VIII, § 7] the following theorem.

THEOREM 10.1. A FUNDAMENTAL SUFFICIENCY THEOREM. *Consider a normal extremaloid E_{12} which is an extremaloid of a field $\mathfrak{F}$. Suppose that the ends of E_{12} satisfy the conditions $\psi_\rho = 0$ and that there is a neighborhood N of these ends in $(x_1y_1x_2y_2)$-space such that no other extremaloid of the field has ends in N satisfying the equations $\psi_\rho = 0$. If at each point of $\mathfrak{F}$ the condition*

$$(10.5) \qquad \mathcal{E}[x, y, p^-(x,y), l^-(x,y), y'] > 0$$

holds for every admissible set (x, y, y') such that $(x, y, y') \neq (x, y, p^-)$ $(x, y, y') \neq (x, y, p^+)$, then the neighborhood N can be so restricted that the inequality $g(C_{34}) > g(E_{12})$ is true for every admissible arc C_{34} in $\mathfrak{F}$ with ends in N satisfying the conditions $\psi_\rho = 0$ and not identically with E_{12}.

Between corner manifolds the inequality (10.5) gives the usual strengthened form of the Weierstrass condition. On the corner manifolds this inequality implies in view of condition (e) above that we have also $\mathcal{E}[x, y, p^+(x,y), l^+(x,y), y'] > 0$ for every admissible set (x, y, y') such that $(x, y, y') \neq (x, y, p^-)$, $(x, y, y') \neq (x, y, p^+)$.

11. Sufficient conditions for a strong relative minimum. The construction of a field of extremaloids for the auxiliary Mayer problem defined in § 10 is embodied in the following theorem.

THEOREM 11.1. *Let E_{12} be a normal extremaloid that is normal on every sub-interval x_1x_3 $(x_1 < x_3 \leq x_2)$, which satisfies conditions III′, IV′, and for which $\Omega_0 \neq 0$ at its corners. Then E_{12} is a member of an n-parameter family of extremaloids*

$$(11.1) \qquad y_i = y_i(x, b_1, \cdots, b_n), \qquad z_i = z_i(x, b_1, \cdots, b_n)$$

whose determinant $|\, y_{ib_j} \,|$ *neither vanishes nor changes sign along* E_{12}; *furthermore,* E_{12} *is an extremaloid of a field* $\mathfrak{F}$ *simply covered by the family.*

The explicit form of such a family may be given as follows. Let $B(b_1, \cdots, b_n)$ be the function

$$B(b) = z_{i2}b_i + (1/2)D_{ij}(b_i - y_{i2})(b_j - y_{j2}),$$

where the constants D_{ij} are defined by equations (9.5) in terms of the end-values of the conjugate system of secondary extremaloids $S_{\theta j}$, U_{ij}, V_{ij} determined in § 9 to satisfy the conditions of Theorem 9.1, and y_{i2}, z_{i2} are the end values at $x = x_{20}$ of the set $y_i(x)$, $z_i(x)$ defining E_{12}. When in equations (3.7) the set (x_0, y_0, z_0) is replaced by the set (x_{20}, b_i, B_{b_i}) an n-parameter family of extremaloids

$$(11.2) \quad y_i = y_i(x, x_{20}, b, B_b) = y_i(x, b), \quad z_i = z_i(x, x_{20}, b, B_b) = z_i(x, b)$$

is defined and contains E_{12} for $b_i = y_{i2}$. The multipliers $\lambda_\alpha(x, b)$ and corner manifolds $\mathfrak{M}_\theta$ associated with this family are determined by equations (3.2) and (3.8). Furthermore, since each extremaloid (11.2) defined by parameter values b_i has on it the element (x_{20}, b_i, B_{b_i}) it is seen that $y_{ib_j} = \delta_{ij}$, $z_{ib_j} = D_{ij}$ at $x = x_{20}$. Hence from Theorem 9.1 it follows that the determinant $|\, y_{ib_j} \,|$ neither vanishes nor changes sign along the extremaloid E_{12} of the family (11.2). This family, therefore, simply covers a neighborhood of E_{12}. In particular, let $\mathfrak{F}$ be a neighborhood of E_{12} which is simply covered by the family, for which the portions of the corner manifolds $\mathfrak{M}_\theta$ of the family which lie in $\mathfrak{F}$ satisfy condition (a) above, and such that if the parameter values of the extremaloid through a point (x, y) of $\mathfrak{F}$ are denoted by $b_i(x, y)$, then the functions

$$(11.3) \quad p_i(x, y) = y_{ix}[x, b(x, y)], \quad l_\alpha(x, y) = \lambda_\alpha[x, b(x, y)]$$

have the sets $[x, y, p(x, y)]$ all admissible. The neighborhood $\mathfrak{F}$ and the functions $p_i(x, y)$, $l_\alpha(x, y)$ defined by (11.3) are seen to satisfy conditions (a)-(e) of § 10. Moreover, on the hyperplane $x = x_{20}$ the Hilbert integral I^* is expressible as

$$I^* = \int F_{y'_i} dy_i = \int B_{b_i} db_i = \int dB$$

and hence is independent of the path. The integral I^* may then be proved to be independent of the path in $\mathfrak{F}$ by an argument that is a slight extension of that used to prove the corresponding result for the case of extremal arcs [see VIII, p. 323; also III, p. 733]. Hence $\mathfrak{F}$ is a field as defined in § 10.

The proof of the following result is the same as for the case of extremal arcs [VIII, p. 323]:

THEOREM 11.2. *Let E_{12} be a normal extremaloid that is normal on x_1x_2, and which is a member of an n-parameter family of extremaloids* (11.1) *whose determinant $|\, y_{ib_j} \,|$ neither vanishes nor changes sign along E_{12}. If the ends of E_{12} satisfy the conditions $\psi_\rho = 0$ $(\rho = 1, \cdots, 2n+1)$ of the auxiliary problem of* § 10, *then there is a neighborhood N of these ends in $(x_1y_1x_2y_2)$-space such that E_{12} is the only extremaloid of the family with ends in N satisfying the conditions $\psi_\rho = 0$.*

The following theorem, which states sufficient conditions for a proper strong relative minimum in the auxiliary Mayer problem formulated in § 10, is seen to be an immediate consequence of the preceding theorems.

THEOREM 11.3. *Let E_{12} be an arc which satisfies the following hypotheses:*

H_1). E_{12} is an admissible arc with r corners and with ends satisfying conditions $\psi_\rho = 0$ $(\rho = 1, \cdots, 2n+1)$ of the auxiliary problem of § 10.

H_2). E_{12} is normal, normal on every sub-interval x_1x_3 of x_1x_2, and satisfies conditions (I), (III'), (IV').

H_3). $\Omega_0 \neq 0$ at the corners on E_{12}.

H_4). The extremaloids of the family (11.1) *determined by Theorem* 11.1 *satisfy condition II'_0 for values of the parameters b_i in a neighborhood of the set $b_i = b_{i0}$ defining E_{12}.*

Then there are neighborhoods $\mathfrak{F}$ of E_{12} in xy-space and N of the ends of E_{12} in $(x_1y_1x_2y_2)$-space such that the inequality $g(C_{34}) > g(E_{12})$ holds for every admissible arc C_{34} in $\mathfrak{F}$ with ends in N satisfying the conditions $\psi_\rho = 0$ and not identical with E_{12}.

The proof of sufficient conditions for a strong relative minimum in the general problem of Mayer follows the methods introduced by Hestenes for the case of extremal arcs. Use is made of a family of fields depending upon $n-1$ parameters which includes for special values of the parameters the field introduced above. In the following discussion it will be assumed that E_{12} satisfies the hypotheses of Theorem 11.1. Let $B_h(b_1, \cdots, b_n)$ $(h = 1, \cdots, n-1)$ be functions of class C'' and such that the determinant $|\, B_{b_i}\, B_{hb_i} \,|$ is different from zero for $b_i = b_{i0}$. Then the $(2n-1)$-parameter family of extremaloids

$$
\begin{aligned}
(11.4) \qquad & y_i = y_i(x, x_{20}, b, B_b + c_hB_{hb}) = y_i(x, b, c), \\
& z_i = z_i(x, x_{20}, b, B_b + c_hB_{hb}) = z_i(x, b, c) \qquad (x_1 \leq x \leq x_2)
\end{aligned}
$$

contains E_{12} for $(x_1, x_2, b, c) = (x_{10}, x_{20}, b_0, 0)$. For this $(2n-1)$-parameter family the determinant

$$\begin{vmatrix} y_{ib_j} & y_{ic_k} & 0 \\ z_{ib_j} & z_{ic_k} & z_i \end{vmatrix}$$

is found to be different from zero for the values $(x, b, c) = (x_{20}, b_0, 0)$.

Since the determinant $|\, y_{ib_j} \,|$ belonging to the family (11.2) neither vanishes nor changes sign along E_{12}, the determinant $|\, y_{ib_j}(x, b, c) \,|$ belonging to the family (11.4) has the same property. Hence the system of equations

$$y_i = y_i(x, b, c) \tag{11.5}$$

has a unique solution

$$b_i = b_i(x, y, c)$$

in a neighborhood $\mathfrak{D}$ of the values (x, y, c) belonging to E_{12}. We shall suppose that $\mathfrak{D}$ is so restricted that for parameter values b_i, c_k sufficiently near to the set b_{i0}, c_{k0} defining E_{12} the portions of the corner manifolds $\mathfrak{M}_\theta$ for the family (11.4) which lie in $\mathfrak{D}$ satisfy condition (a) of § 10; moreover, if

$$\begin{aligned} p_i(x, y, c) &= y_{ix}[x, b(x, y, c), c], \\ l_a(x, y, c) &= \lambda_a[x, b(x, y, c), c], \end{aligned} \tag{11.6}$$

then the sets $[x, y, p(x, y, c)]$ are all admissible.

By the same argument as used above we have that on the hyper-plane $x = x_{20}$ in xy-space the Hilbert integral I^* is independent of the path for fixed values of the parameters c_k. It follows that for each set c_k the region $\mathfrak{F}$ of points (x, y) whose elements are all in $\mathfrak{D}$ forms a field with slope functions and multipliers defined by equations (11.6). If in each of these fields the Weierstrass $\mathcal{E}$-function with arguments p_i, l_a defined by (11.6) is positive for every admissible set (x, y, y') distinct from (x, y, p^+) and (x, y, p^-), then the methods used by Hestenes [IX] for the case of extremal arcs apply to give sufficient conditions for a proper strong relative minimum. We have the following theorem:

THEOREM 11.4. *Let E_{12} be an arc which satisfies the hypotheses (H_1), (H_2), (H_3) of Theorem* 11.3 *and the following strengthened form of hypothesis (H_4) of that theorem.*

H^*_4). The extremaloids of the family (11.4) *satisfy condition II'_0 for values of the parameters b_i, c_k in a neighborhood of the values $b_i = b_{i0}$, $c_k = c_{k0}$ defining E_{12}.*

Then there are neighborhoods $\mathfrak{F}$ of E_{12} in xy-space and N of the ends of E_{12} in $(x_1y_1x_2y_2)$-space such that the inequality $g(C_{34}) > g(E_{12})$ holds for every admissible arc C_{34} in $\mathfrak{F}$ with ends in N satisfying the conditions $\psi_\mu = 0$ and not identical with E_{12}.

Suppose now that the functions ψ_μ are continuous at every pair of distinct or coincident points in a neighborhood of those belonging to E_{12}. If the ends of E_{12} are the only pair of distinct or coincident points on E_{12} satisfying the conditions $\psi_\mu = 0$, then for every $\delta > 0$ there exists a neighborhood $\mathfrak{F}$ of E_{12} such that every pair of points in $\mathfrak{F}$ satisfying the equations $\psi_\mu = 0$ necessarily has one of the points of the pair in each of the two δ-neighborhoods of the ends of E_{12}. Hence by suitably restricting the neighborhood $\mathfrak{F}$ of E_{12} we obtain the following corollary.

COROLLARY. *Let E_{12} be an arc satisfying the hypotheses of Theorem* 11.4. *If further the ends of E_{12} are the only pair of distinct or coincident points on E_{12} satisfying the conditions $\psi_\mu = 0$, then there is a neighborhood $\mathfrak{F}$ of E_{12} in xy-space such that the inequality $g(C_{34}) > g(E_{12})$ holds for every admissible arc C_{34} in $\mathfrak{F}$ with ends satisfying the conditions $\psi_\mu = 0$ and not identical with E_{12}.*

In conclusion, it is to be remarked that in Theorems 11.3 and 11.4 the condition II'_0 has been assumed to be satisfied along the members of certain particular families of extremaloids. Let E denote an arbitrary extremaloid with equations (2.1) which has a corner for $x = t$, and for which $\Omega_0 < 0$ at this corner. The results of § 4 imply that there are sets of elements (x, y, y', Y', λ) for which the set $[x, y, y', \lambda]$ is in a given neighborhood of either the set $[t, y(t), y'(t-), \lambda(t-)]$ or the set

$$[t, y(t), y'(t+), \lambda(t+)],$$

and for which $\mathcal{E}[x, y, y', Y', \lambda] < 0$. Consequently, for the non-parametric problem it seems necessary to phrase the strengthened Weierstrass condition in the manner indicated above. In this respect, therefore, the statement of sufficient conditions for a strong relative minimum is more cumbersome for non-parametric problems than for the corresponding parametric problems [see IV, § 5].

Remarks. [*Added to proof sheets, November 24, 1934.*]

In a recent paper Hestenes has proved sufficiency theorems for an extremal arc in the problem of Bolza assuming no normality conditions aside from the existence of multipliers of the form $\lambda_0 = 1$, $\lambda_\alpha(x)$, by modifying the usual form of the Mayer condition. Hestenes' methods may be extended to establish corresponding theorems for the problem of discontinuous solutions here studied.* Indeed, such sufficiency theorems may be established by the same

* *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 793-818.

method as used above. This fact is a consequence of a result that has recently been established independently by Morse and the author.*

BIBLIOGRAPHY

I. Bliss, "The problem of Mayer with variable end points," *Transactions of the American Mathematical Society*, vol. 19 (1918), pp. 305-314.

II. Cope, "An analogue of Jacobi's condition for the problem of Mayer with variable end points," Dissertation, The University of Chicago (1927).

III. Bliss, "The problem of Lagrange in the calculus of variations," *American Journal of Mathematics*, vol. 52 (1930), pp. 673-744.

IV. Graves, "Discontinuous solutions in space problems of the calculus of variations," *American Journal of Mathematics*, vol. 52 (1930), pp. 1-28.

V. Graves, "Discontinuous solutions in the calculus of variations," *Bulletin of the American Mathematical Society*, vol. 36 (1930), pp. 831-846.

VI. Hefner, "The condition of Mayer for discontinuous solutions of the Lagrange problem," *Contributions to the Calculus of Variations 1931-1932*—Theses submitted to the Department of Mathematics of the University of Chicago, pp. 95-130.

VII. Reid, "Boundary value problems associated with the calculus of variations," *American Journal of Mathematics*, vol. 54 (1932), pp. 769-790.

VIII. Bliss and Hestenes, "Sufficient conditions for a problem of Mayer in the calculus of variations," *Transactions of the American Mathematical Society*, vol. 35 (1933), pp. 305-326.

IX. Hestenes, "Sufficient conditions for the general problem of Mayer with variable end points," *Transactions of the American Mathematical Society*, vol. 35 (1933), pp. 479-490.

X. Reid, "Analogues of the Jacobi condition for the problem of Mayer in the calculus of variations," *Annals of Mathematics*, vol. 35 (1934), pp. 836-848.

UNIVERSITY OF CHICAGO,
CHICAGO, ILLINOIS.

* For abstracts of results, see *Bulletin of the American Mathematical Society*, vol. 40 (1934), p. 665 and p. 666. The central theorem concerning secondary extremal arcs which is involved has also been established by Hestenes.

ON THE INVERSION FORMULA FOR FOURIER-STIELTJES TRANSFORMS IN MORE THAN ONE DIMENSION.

By E. K. HAVILAND.

An inversion formula for Fourier-Stieltjes transforms in one dimension, given by P. Lévy,* has been extended to more than one dimension by V. Romanovsky † in a formal treatment. It is the purpose of the present paper to give an exact proof of this formula. For simplicity, the proof is given for the case of two dimensions, but its extension to n-dimensions is given by precisely the same methods.

It is known ‡ that if a function $f(x, y)$ is

(i) bounded throughout the entire (xy)-plane,
(ii) of bounded variation in every finite region of the plane,
(iii) absolutely integrable over the entire (xy)-plane,

then to the formula

$$(1) \qquad E(s,t) = (2\pi)^{-2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) \exp\left[i(sx+ty)\right] dxdy$$

there corresponds the inverse formula:

$$(2) \qquad f(u,v) = \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \right) E(s,t) \exp\left[-i(us+vt)\right] dsdt,$$

where the latter double integral is to be considered as a Cauchy principal value, as indicated by the parenthesis, and (u, v) is any continuity point of $f(u, v)$.

Accordingly, we proceed to show that if $\phi(E)$ is a distribution function,§ and $F(x, y)$ is the corresponding point function,¶ then ‖

---

* P. Lévy, *op. cit.*, pp. 166-167. References are collected at the end of the paper.

† V. Romanovsky, *loc. cit.*, pp. 36-40.

‡ Cf., e. g., V. Romanovsky, *loc. cit.*, pp. 36-38; also S. Bochner, *op. cit.*, p. 202, Theorem 66. These authors do not give the theorem precisely in the above form, but a proof of the latter may be obtained by the use of their methods together with certain analogues of the Second Theorem of the Mean given by W. H. Young, *loc. cit.*, p. 290, equations (23), (24), (25) and (26).

§ The monotone absolutely additive set function $\phi(E)$ is said to be a distribution function if $0 \leqq \phi(E) \leqq 1$ and $\phi(S) = 1$, where S denotes the whole (xy)-plane. Cf. E. K. Haviland, *loc. cit.*, p. 627.

¶ Cf. J. Radon, *loc. cit.*, pp. 1304-1305.

‖ $R + P_{\xi\eta}$ represents the vector addition of the quadrant R and the point $P_{\xi\eta} = (\xi, \eta)$.

$$(3)\quad g(x,y) = G(x,y;\xi,\eta)$$
$$= \phi(R + P_{\xi\eta}) - \phi(R + P_{\xi 0}) - \phi(R + P_{0\eta}) + \phi(R)$$
$$= F(x+\xi, y+\eta) - F(x+\xi, y) - F(x, y+\eta) + F(x,y),$$

where R: $(-\infty < u < x; -\infty < v < y)$, as a function of x and y, fulfills conditions (i), (ii), (iii) above.

That $G(x,y;\xi,\eta)$ is bounded for all (x,y), ξ and η being fixed, is seen at once from the fact that ϕ is a distribution function, in which case $|G| < 4$.

That $G(x,y;\xi,\eta)$ is of bounded variation with respect to (x,y) may be seen by examining its four terms separately. In the first place, $F(x+\xi, y+\eta)$ is a monotone function of (x,y); for if $x_2 > x_1$, $y_2 > y_1$, and R_{ij}: $(-\infty < x < x_i; -\infty < y < y_j)$,

$$F(x_2+\xi, y_2+\eta) - F(x_2+\xi, y_1+\eta) - F(x_1+\xi, y_2+\eta) + F(x_1+\xi, y_1+\eta)$$
$$= \phi(R_{22} + P_{\xi\eta}) - \phi(R_{21} + P_{\xi\eta}) - \phi(R_{12} + P_{\xi\eta}) + \phi(R_{11} + P_{\xi\eta})$$
$$= \phi(E_1 + P_{\xi\eta}) \geqq 0, \text{ where } E_1\text{: } (x_1 \leqq x < x_2;\ y_1 \leqq y < y_2),$$

since ϕ is a distribution function. Similarly, $F(x+\xi, y)$, $F(x, y+\eta)$ and $F(x,y)$ are bounded monotone functions of (x,y), so that $G(x,y;\xi,\eta)$ is of bounded total variation throughout the (x,y)-plane.*

To prove $G(x,y;\xi,\eta)$ absolutely integrable throughout the (x,y)-plane for fixed ξ and η, we observe that †

$$\int_a^b\int_c^d |\, G(x,y;\xi,\eta)\,|\, dxdy = \pm \int_a^b\int_c^d G(x,y;\xi,\eta)\,dxdy$$
$$= \pm \int_a^b\int_c^d [F(x+\xi, y+\eta) - F(x+\xi, y) - F(x, y+\eta) + F(x,y)]\,dxdy$$
$$= \pm \left\{ \int_{a+\xi}^{b+\xi}\int_{c+\eta}^{d+\eta} F(x,y)\,dxdy - \int_{a+\xi}^{b+\xi}\int_c^d F(x,y)\,dxdy \right.$$
$$\left. - \int_a^b\int_{c+\eta}^{d+\eta} F(x,y)\,dxdy + \int_a^b\int_c^d F(x,y)\,dxdy \right\}$$
$$= \pm \left\{ \int_{a+\xi}^{b+\xi} \left[\int_{c+\eta}^{d+\eta} - \int_c^d \right] F(x,y)\,dxdy - \int_a^b \left[\int_{c+\eta}^{d+\eta} - \int_c^d \right] F(x,y)\,dxdy \right\}$$
$$= \pm \left\{ \int_{a+\xi}^{b+\xi} \left[\int_d^{d+\eta} - \int_c^{c+\eta} \right] F(x,y)\,dxdy \right.$$
$$\left. - \int_a^b \left[\int_d^{d+\eta} - \int_c^{c+\eta} \right] F(x,y)\,dxdy \right\}$$

* A statement by B. H. Camp, *loc. cit.*, p. 45 that a bounded monotone function is not necessarily of bounded variation in the sense of Hardy or Radon is incorrect.

† That the double and iterated integrals given below exist and are equivalent may be seen by consulting, e. g., E. W. Hobson, *op. cit.*, pp. 481 and 483.

$$= \pm \left\{ \int_{a+\xi}^{b+\xi} \int_0^{\eta} [F(x, y+d) - F(x, y+c)]\, dx dy \right.$$
$$\left. - \int_a^b \int_0^{\eta} [F(x, y+d) - F(x, y+c)]\, dx dy \right\}$$
$$= \pm \left\{ \left[\int_{a+\xi}^{b+\xi} - \int_a \right] \int_0^{\eta} [F(x, y+d) - F(x, y+c)]\, dx dy \right\}$$
$$= \pm \left\{ \left[\int_b^{b+\xi} - \int_a^{a+\xi} \right] \int_0^{\eta} [F(x, y+d) - F(x, y+c)]\, dx dy \right\}$$
$$= \pm \int_0^{\xi} \int_0^{\eta} [F(x+b, y+d) - F(x+b, y+c)$$
$$- F(x+a, y+d) + F(x+a, y+c)]\, dx dy.$$

But as $F(x, y)$ is the point function associated with the distribution function $\phi(E)$, the integrand $\to 0$ uniformly with respect to (x, y) in $(0 \leqq x \leqq \xi;\ 0 \leqq y \leqq \eta)$ as $a, b \to +\infty$ or $\to -\infty$, c and d being arbitrary; or as $c, d \to +\infty$ or $\to -\infty$, a and b being arbitrary. Hence

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} | G(x, y; \xi, \eta) |\, dx dy < +\infty, \qquad \text{q. e. d.}$$

The Fourier-Stieltjes transform $\Lambda(s, t; \phi)$ of the distribution function $\phi(E)$ is given by *

$$\Lambda(s, t; \phi) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp [i(sx + ty)]\, d_{xy}\phi(E)$$
$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp [i(sx + ty)]\, d_{xy}F(x, y),$$

where $F(x, y)$ is the point function corresponding to the absolutely additive set function ϕ and s and t are real. It follows that if $s \neq 0$, $t \neq 0$,

$$\Lambda(s, t; \phi)\,(is)^{-1}(it)^{-1}\,[\exp(-is\xi) - 1]\,[\exp(-it\eta) - 1]$$
$$= -(st)^{-1} \left\{ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp [i\{s(x-\xi) + t(y-\eta)\}]\, d_{xy}\phi(E) \right.$$
$$- \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp [i\{s(x-\xi) + ty\}]\, d_{xy}\phi(E)$$
$$\left. - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp [i\{sx + t(y-\eta)\}]\, d_{xy}\phi(E) \right.$$

* The integral is the limit as the rectangle $R \to S$, the entire (x, y)-plane, of $\int\int_R \exp[i(sx + ty)]\, d_{xy}\phi(E)$, the latter being defined by J. Radon, *loc. cit.*, pp. 1322-1324. The second and third members of this equation represent merely different notations for the same integral.

$$+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left[i(sx+ty)\right]d_{xy}\phi(E)\Big\}$$

$$=-(st)^{-1}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left[i(sx+ty)\right]d_{xy}[\phi(E+P_{\xi\eta})$$
$$-\phi(E+P_{\xi 0})-\phi(E+P_{0\eta})+\phi(E)],$$

so

$$(4)\quad \Lambda(s,t;\phi)(is)^{-1}(it)^{-1}\left[\exp\left(-is\xi\right)-1\right]\left[\exp\left(-it\eta\right)-1\right]$$
$$=-(st)^{-1}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left[i(sx+ty)\right]d_{xy}g(x,y)$$

by the preceding footnote and equation (3). If we set

$$(5)\qquad \exp\left[i(sx+ty)\right]=f(x,y),$$

we obtain by the partial integration formula of W. H. Young:*

$$(6)\quad (st)^{-1}\int_a^b\int_c^d g(x,y)\,d_{xy}f(x,y)-(st)^{-1}\int_a^b\int_c^d f(x,y)\,d_{xy}g(x,y)$$
$$=(st)^{-1}\Big\{\int_a^b f(x,c)\,d_x g(x,c)+\int_c^d f(a,y)\,d_y g(a,y)$$
$$-\int_a^b f(x,d)\,d_x g(x,d)-\int_c^d f(b,y)\,d_y g(b,y)$$
$$+[f(b,d)g(b,d)-f(b,c)g(b,c)-f(a,d)g(a,d)+f(a,c)g(a,c)]\Big\}.$$

We shall show that as $a, c\to-\infty$ and $b, d\to+\infty$, all terms on right of the preceding equation vanish. By the one-dimensional partial integration formula:

$$(st)^{-1}\int_a^b f(x,c)\,d_x g(x,c)=-(st)^{-1}\int_a^b g(x,c)\,d_x f(x,c)$$
$$+(st)^{-1}[g(x,c)f(x,c)]_{x=a}^{x=b}.$$

Now from (3) it is seen that

$$(7)\qquad g(x,y)=G(x,y;\xi,\eta)=\pm\phi(Q+P_{xy}),$$

where Q: $(0\leqq u<\xi;\ 0\leqq v<\eta)$ or $(\xi\leqq u<0;\ \eta\leqq v<0)$ or $(0\leqq u<\xi;\ \eta\leqq v<0)$ or $(\xi\leqq u<0;\ 0\leqq v<\eta)$, depending on the signs of ξ and η. In the two former cases, $g(x,y)$ is non-negative; in the two latter, it is non-positive. Since ϕ is a distribution function, there exists a rectangle J so large that $0\leqq\phi(S-J)<\epsilon/4$, where S is the entire (x,y)-plane. Accordingly, if (x,y) is so distant that $Q+P_{xy}$ lies outside J,

* Cf. W. H. Young, *loc. cit.*, p. 282.

$0 \leqq \phi(Q + P_{xy}) < \epsilon/4$; from which fact it follows, as $|f(x,y)| = 1$, that there is an $M_1 > 0$ so large that if $|a|, |b|, |c|, |d| > M_1$,

$$\begin{aligned} &|st|^{-1} \left| [g(x,c)f(x,c)]_{x=a}^{x=b} \right| \\ &= |st|^{-1} |G(b,c;\xi,\eta)f(b,c) - G(a,c;\xi,\eta)f(a,c)| < \epsilon/2. \end{aligned}$$

Again,

$$\begin{aligned} &|(st)^{-1}\int_a^b g(x,c)\,d_x f(x,c)| \leqq |t|^{-1}\int_a^b |G(x,c;\xi,\eta)|\,dx \\ &= \pm t^{-1}\int_a^b [F(x+\xi,c+\eta) - F(x+\xi,c) - F(x,c+\eta) + F(x,c)]dx \\ &= \pm t^{-1}\left[\int_{a+\xi}^{b+\xi} - \int_a^b\right][F(x,c+\eta) - F(x,c)]dx \\ &= \pm t^{-1}\left[\int_b^{b+\xi} - \int_a^{a+\xi}\right][F(x,c+\eta) - F(x,c)]dx \\ &= \pm t^{-1}\int_0^{\xi} [F(b+x,c+\eta) - F(b+x,c) \\ &\qquad - F(a+x,c+\eta) + F(a+x,c)]dx. \end{aligned}$$

Thus, as t and η are fixed, the integrand can, for all x in $[0,\xi]$ be made arbitrarily small by taking $|c|$ sufficiently large. Consequently, as $t \neq 0$,

$$|(st)^{-1}\int_a^b g(x,c)\,d_x f(x,c)| < \epsilon/2 \text{ if } |c| > M_2, \text{ say.}$$

Therefore if $|a|, |b|, |c|, |d|$ are all $> \max(M_1, M_2)$,

$$|(st)^{-1}\int_a^b f(x,c)\,d_x g(x,c)| < \epsilon.$$

Precisely similar reasoning shows that

$$|(st)^{-1}\int_a^b f(x,d)\,d_x g(x,d)| < \epsilon, \; |(st)^{-1}\int_c^d f(a,y)\,d_y g(a,y)| < \epsilon$$

and $|(st)^{-1}\int_c^d f(b,y)\,d_y g(b,y)| < \epsilon$ if $|a|, |b|, |c|, |d|$ are sufficiently large. Finally, for fixed $s, t \neq 0$,

$$\begin{aligned} &|(st)^{-1}\{f(b,d)g(b,d) - f(b,c)g(b,c) - f(a,d)g(a,d) + f(a,c)g(a,c)\}| \\ &\qquad \leqq |st|^{-1}\{|g(b,d)| + |g(b,c)| + |g(a,d)| + |g(a,c)|\}. \end{aligned}$$

From equation (7) it can be seen that the foregoing expression may be made arbitrarily small by taking $|a|, |b|, |c|, |d|$ sufficiently large. Hence for all fixed $s, t \neq 0$, the right-hand side of (6) vanishes as $a, c \to -\infty$; $b, d \to +\infty$. The left-hand side then becomes

$$(st)^{-1}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\,d_{xy}f(x,y) - (st)^{-1}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,d_{xy}g(x,y),$$

or by virtue of (4) and (5):

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\,\exp[i(sx+ty)]\,dxdy + (is)^{-1}(it)^{-1}\Lambda(s,t;\phi)\,[\exp(-is\xi)-1]\,[\exp(-it\eta)-1].$$

As the limit as $s\to 0$ and/or $t\to 0$ of the left-hand side of (6) exists, the same is true of the right-hand side, so that we have for any fixed s, t:

$$(2\pi)^{-2}(is)^{-1}(it)^{-1}\Lambda(s,t;\phi)\,[\exp(-is\xi)-1]\,[\exp(-it\eta)-1 = (2\pi)^{-2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\,\exp[i(sx+ty)]\,dxdy.$$

Then by (2):

$$(8)\quad g(u,v) = (2\pi)^{-2}\left(\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\right)(is)^{-1}(it)^{-1}\Lambda(s,t;\phi)\,[\exp(-is\xi)-1]\,[\exp(-it\eta)-1]\,\exp[-i(us+vt)]\,dsdt,$$

where $g(u,v) = G(u,v;\xi,\eta) = \phi(Q+P_{uv})$. This is the required inversion formula. It has been proved under the assumption that (u,v) is a continuity point of g, and the latter will certainly be the case if $Q+P_{uv}$ is a non-singular * rectangle of ϕ. Moreover, it is known that the non-singular rectangles of ϕ are everywhere dense, so that ϕ is essentially determined by the inversion formula.*

With the help of (8) and the multiplication rule for Fourier-Stieltjes transforms,† one can prove the Continuity Theorem for Fourier-Stieltjes transforms: ‡ If $\{\Lambda_n(s,t) = \Lambda(s,t;\phi_n)\}$ is a sequence of Fourier-Stieltjes transforms of distribution functions such that the sequence converges uniformly in every finite rectangle of the st-plane to a function $\Lambda(s,t)$, then $\Lambda(s,t)$ is the Fourier-Stieltjes transform of a distribution function ϕ and $\phi_n\to\phi$ on all non-singular rectangles of the latter.

---

* Cf. E. K. Haviland, *loc. cit.*, p. 628.

† For a proof of this rule, cf. E. K. Haviland, *loc. cit.*, p. 651, Theorem V. In the proof of this theorem, it is to be noted that the last line on p. 652 should read

"$$\Big|\sum_{k=1}^{N}\{\psi(R_k) - \sum_{j=1}^{m}\phi_1(R_k - P_j)\phi_2(E_j)\}\Big| < \epsilon/2$$"

and correspondingly the summation with respect to k at the top of p. 653 should be taken beneath the absolute value sign. It is to be noted that $\sum_{j=1}^{m} E_j$ need not extend beyond J.

‡ The one-dimensional case of this theorem is treated by P. Lévy, *op. cit.*, the two-dimensional case by V. Romanovsky, *loc. cit.*, p. 41 and by S. Bochner, loc. cit., II, p. 403.

Finally, if $\Lambda(s, t; \phi)$ is absolutely integrable, it is possible to differentiate with respect to ξ and η beneath the integral sign in (8) and the resulting second mixed derivative is a continuous function of ξ and η in any finite region of the $(\xi\eta)$-plane. As

$$\phi(Q + P_{uv}) = F(u + \xi, v + \eta) - F(u + \xi, v) - F(u, v + \eta) + F(u, v),$$

we may then infer, on setting $u = v = 0$, that $\phi(E)$ is absolutely continuous and that the point function $F(x, y)$ corresponding to it posseses a continuous mixed derivative $\partial^2 F/\partial x \partial y$.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

S. Bochner, I.: *Vorlesungen ueber Fouriersche Integrale* (Leipzig, 1932); II. "Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse." *Mathematische Annalen*, vol. 108 (1933), pp. 378-410.

B. F. Camp, "Singular multiple integrals with applications to series," *Transactions of the American Mathematical Society*, vol. 14 (1913), pp. 42-64.

E. K. Haviland, "On the theory of absolutely additive distribution functions," *American Journal of Mathematics*, vol. 56 (1934), pp. 625-658.

E. W. Hobson, *Theory of functions of a real variable*, vol. 1 (Cambridge, 1921).

P. Lévy, *Calcul des probabilités* (Paris, 1925).

J. Radon, "Theorie und Anwendungen der absolut additiven Mengenfunktionen," *Sitzungsberichte der mathematischen-naturwissenschaftlichen Klasse der Kaiserl. Akademie zu Wien*, vol. 122 (1913), pp. 1295-1438.

V. Romanovsky, "Sur un théorème limite du calcul des probabilités," *Recueil Mathématique de la Société Mathématique de Moscou*, vol. 36 (1929), pp. 36-64.

W. H. Young, "On multiple integration by parts and the second theorem of the mean," *Proceedings of the London Mathematical Society*, ser. 2, vol. 16 (1917), pp. 273-293.

A NOTE ON THE DISTRIBUTION OF THE ZEROS OF THE ZETA-FUNCTION.

By AUREL WINTNER.

The results of Littlewood * on the remainder term of the prime number theorem depend upon a close connection between the distribution of the prime numbers on the one hand and the *cyclical* distribution of the sequence

$$\gamma_1, \gamma_2, \cdots \qquad (0 < \gamma_1 \leqq \gamma_2 \leqq \cdots) \tag{1}$$

on the other hand, where $\rho_1 = \beta_1 + i\gamma_1$, $\rho_2 = \beta_2 + i\gamma_2, \cdots$ denote the zeros of $\zeta(s)$ in the upper half-plane. While for the purposes of Littlewood only a consequence of the Dirichlet approximation theorem is relevant, it seems to be worth while to determine the asymptotic cyclical distribution of (1). The object of the present note is to point out the fact that *the sequence* (1) *is asymptotically equidistributed* † *to modulus c*, where c is any non-vanishing real number. In order to simplify the formulae we shall choose $c = 1$; it will be clear that the proof holds for any c.

We have to prove that

$$K(n; x)/n \to x \qquad (n \to \infty), \tag{2}$$

where x is any point of the interval $0 \leqq x < 1$ and $K(n; x)$ denotes the number of values k satisfying both inequalities

$$k \leqq n, \qquad \gamma_k \leqq x. \tag{3}$$

The truth of (2) would follow in virtue of the Riemann-Mangoldt asymptotic formula from a known general lemma ‡ if we knew the sequence of the differences $\gamma_{k+1} - \gamma_k$ to be a monotone sequence. In reality, we do not even know that $\gamma_{k+1} - \gamma_k = 0$ is impossible, and the numerical results of Gram § show that the sequence of the differences $\gamma_{k+1} - \gamma_k$ is *not* monotone, at least

* Cf. A. E. Ingham, "The distribution of prime numbers," *Cambridge Tracts*, No. 30, London, 1932, pp. 96-104.

† "Gleichverteilt" in the sense of Weyl. Cf. G. Pólya and G. Szegö, *Aufgaben und Lehrsätze*, Berlin (1925), vol. I, pp. 70-73.

‡ *Ibid.*, p. 72.

§ J. P. Gram, "Note sur les zéros de la fonction $\zeta(s)$ de Riemann," *Acta Mathematica*, vol. 27 (1903), p. 297.

if small values of k are not discarded. For the proof of (2) one does not need, however, anything but the Riemann-Mangoldt asymptotic formula in its roughest form, viz.

$$N(T) \sim L(T/2\pi) \qquad (T \to \infty), \tag{4}$$

where

$$L(T) = T \log T \tag{5}$$

and $N(T)$ denotes the number of values k satisfying the inequality

$$\gamma_k \leqq T. \tag{6}$$

By the definitions (3) and (6) of $K(n;x)$ and $N(T)$ we have the identity

$$\begin{aligned} K(N(T);x) = {} & N(1+x) - N(1) + N(2+x) - N(2) + \cdots \\ & + N([T]-1+x) - N([T]-1) + N(\min\{T, [T]+x\}) - N([T]) \end{aligned}$$

so that

$$K(N(T);x) = \sum_{k=1}^{[T]} \{N(k+x) - N(k)\} + o(N(T)) \tag{7}$$

in virtue of (4) and (5). Now (4) implies by a well known limit theorem that

$$\sum_{k=1}^{[T]} \{N(k+x) - N(k)\} \sim \sum_{k=1}^{[T]} \{L((k+x)/2\pi) - L(k/2\pi)\}$$

for every fixed x. This is easily seen to be

$$\sim \int_1^T \{L((x+k)/2\pi) - L(k/2\pi)\}\, dk,$$

hence $\sim xL(T/2\pi)$ in virtue of l'Hôspital's rule. Thus the sum $\sum_{k=1}^{[T]}$ occurring in (7) is $\sim xL(T/2\pi)$, hence $= xN(T) + o(N(T))$ in virtue of (4). Consequently $K(N(T);x) \sim xN(T)$, which is the same thing as (2).

THE JOHNS HOPKINS UNIVERSITY.

PARALLELISM AND EQUIDISTANCE IN RIEMANNIAN GEOMETRY.

By R. M. PETERS.

It is the purpose of this paper to generalize to Riemannian spaces the concepts and some of the resuts given by Graustein * for classical differential geometry. The angular spread, measure of the deviation from parallelism, in the sense of Levi-Civita, of one of two families of curves on a two-dimensional surface with respect to the other, becomes in Riemannian space the associate curvature of one congruence of curves with respect to another. The concept of distantial spread as a measure of the deviation from equidistance of the one family of curves with respect to the other suggests two corresponding spreads in Riemannian space: first, the distantial spread of a family of hypersurfaces with respect to a congruence of transversals; secondly, the distantial spread of one congruence of curves with respect to another. The latter concept is developed here only for the case when the curves of the two congruences lie on two-dimensional surfaces.

The paper deals with the relations between associate curvatures and distantial spreads and the conclusions which they imply concerning the connection between parallelism and equidistance.

It is assumed throughout that the linear element of the n-dimensional space V_n is defined by the positive definite quadratic form $ds^2 = g_{ij}dx^i dx^j$, the x's being the coördinates of the V_n, and the g_{ij}'s real analytic functions of the x's. The curves and surfaces considered are always assumed to be real and analytic.

1. *Associate curvature.* If C is a curve in V_n represented parametrically in terms of its arc by the equations $x^i = x^i(s)$, and λ^i are the components of a unit vector defined at each point of C, the associate curvature of the vector λ^i with respect to the curve C is

$$1/r = (g_{ij}\mu^i\mu^j)^{\frac{1}{2}} \tag{1}$$

* Graustein, "Parallelism and Equidistance in Classical Differential Geometry," *Transactions of the American Mathematical Society*, vol. 34 (1932), pp. 557-593.

where

$$\mu^i = (dx^k/ds)\lambda^i{}_{,k} \tag{2}$$

$\lambda^i{}_{,k}$ being the components of the covariant derivative of the vector λ^i with respect to the fundamental tensor g_{ij}.

The associate curvature $1/r$ is a measure of the deviation of the vectors λ^i from parallelism with respect to C in that its vanishing is the condition for parallelism.

The vector μ^i is known as the associate curvature vector of the vector λ^i with respect to the curve C; it is orthogonal to the vector λ^i.

If the vectors λ^i are, in particular, tangent to C, μ^i and $1/r$ become respectively the principal normal vector and (first) curvature of C.

That the associate curvature of the vector λ^i with respect to C is based essentially on angle is evident from the fact that it can also be defined as

$$1/r = \lim_{\Delta s \to 0} \Delta\theta/\Delta s \tag{3}$$

where $\Delta\theta$ is the angle between the vector λ^i at a point $P' : (s + \Delta s)$ of C and the direction at P' parallel with respect to C to the vector λ^i at $P : (s)$. The possibility of this definition was suggested by Bianchi * and carried through by Lipka † for the special case of the curvature of C.

Instead of the single curve C, we shall consider ordinarily n linearly independent congruences of curves C_k, $(k = 1, 2, \cdots, n)$, with unit tangent vectors $\lambda_{k|}{}^i = dx^i/ds_k$, s_k being the arc of the curves C_k. Then the associate curvature vector of the vectors $\lambda_{h|}{}^i$ (tangent to the curves C_h) with respect to the curves C_k is

$$\mu_{hk|}{}^j = \lambda_{k|}{}^i \lambda_{h|}{}^j{}_{,i}. \tag{4}$$

We shall call this vector *the associate curvature vector of the curves C_h with respect to the curves C_k*, and its length

$$1/r_{hk} = (g_{ij}\mu_{hk|}{}^i \mu_{hk|}{}^j)^{1/2} \tag{5}$$

the associate curvature of the curves C_h with respect to the curves C_k.

2. *Equidistance of families of hypersurfaces with respect to congruences*

---

* Bianchi, "Sul parallelismo vincolato di Levi-Civita nella metrica degli spazi curvi," *Rendiconti della Reale Accademia di Napoli*, ser. 3a, vol. 28 (1922), p. 161.

† Lipka, "Sulla curvatura geodetica delle linee appartenenti ad una varieta qualunque," *Rendiconti della Reale Accademia dei Lincei*, ser. 5, vol. 31 (1922), pp. 353-356.

of curves. Let $f_k(x^1, x^2, \cdots, x^n) = c_k$, $(k = 1, 2, \cdots, n)$, be the equations of n independent families of hypersurfaces S_k in V_n. These hypersurfaces intersect in n linearly independent congruences of curves C_h, $(h = 1, 2, \cdots, n)$, of which $\lambda_{h|}{}^i = dx^i/ds_h$ are the unit tangent vectors. The hypersurfaces S_k of the family $f_k = c_k$ contain the $n-1$ congruences C_h, $(h = 1, 2, \cdots, n;\ h \neq k)$, the curves C_k being transversals of these hypersurfaces.

Let d_k be the distance, measured along an arbitrary curve C_k, between a hypersurface $S_k : f_k = f_k{}^0$ and a neighboring hypersurface $S_k : f_k = f_k{}^0 + \Delta f_k$, and take the logarithmic directional derivative of d_k in the positive direction of a curve C_h lying in the hypersurface $S_k : f_k = f_k{}^0$. The limit of this derivative when Δf_k approaches zero is defined as *the distantial spread of the hypersurfaces S_k with respect to the curves C_k, measured along the curves C_h*:

$$1/b_{hk} = \lim_{\Delta f_k \to 0} \partial \log d_k/\partial s_h \qquad (h \neq k). \tag{6}$$

If s_k is the directed arc of an individual curve C_k,

$$ds_k = (\partial f_k/\partial s_k)^{-1} df_k, \tag{7}$$

where $\partial f_k/\partial s_k$ is the directional derivative of f_k in the positive direction of the curves C_k. Hence, we obtain as the value of the foregoing limit

$$1/b_{hk} = -\partial \log |\, \partial f_k/\partial s_k \,|\, /\partial s_h, \qquad (h \neq k). \tag{8}$$

If $1/b_{hk} \equiv 0$, $(h = 1, 2, \cdots, n;\ h \neq k)$, then $\partial f_k/\partial s_k$ is a constant along each curve C_h, $(h \neq k)$, and hence is a function of f_k. Then, ds_k, as given by (7) is an exact differential of a function $F(f_k)$, and $s_k = F(f_k)$ is the common directed arc, measured from a fixed hypersurface S_k, of all the curves C_k.

Noting that this argument may be reversed, we conclude that *the distantial spreads $1/b_{hk}$ of the hypersurfaces S_k with respect to the curves C_k in the directions of the curves C_h, $h \neq k$, all vanish if and only if the hypersurfaces S_k are equidistant with respect to the curves C_k,* in that each two of them cut segments of equal length from all the curves C_k. Thus the quantities $1/b_{hk}$ are actually measures of the deviation from equidistance of the hypersurfaces S_k with respect to the curves C_k.

Let θ_k denote the angle between the normal at a point P to the hypersurface S_k through P and the direction of the curve C_k through P. Since $\partial f_k/\partial s_k = (\Delta_1 f_k)^{1/2} \cos\theta_k$, where $(\Delta_1 f_k)^{1/2} = (g^{ij} f_{k|,i} f_{k|,j})^{1/2}$ is the magnitude of the gradient of f_k, we have

$$\begin{aligned} 1/b_{hk} &= -(\Delta_1 f_k)^{-1/2} \sec\theta_k\, \partial (f_{k|,i} \lambda_{k|}{}^i)/\partial s_h \\ &= -(\Delta_1 f_k)^{-1/2} \sec\theta_k (f_{k|,i} \lambda_{k|}{}^i)_{,j} \lambda_{h|}{}^j \\ &= -(\Delta_1 f_k)^{-1/2} \sec\theta_k\, f_{k|,ij} \lambda_{k|}{}^i \lambda_{h|}{}^j + f_{k|,i} \lambda_{k|}{}^i{}_{,j} \lambda_{h|}{}^j. \end{aligned}$$

Now

$$f_{k|,ij}\lambda_{k|}{}^{i}\lambda_{h|}{}^{j} = \partial f_{k|,j}\lambda_{h|}{}^{j}/\partial s_{k} - f_{k|,j}\lambda_{h|}{}^{j}{}_{,i}\lambda_{k|}{}^{i}$$

since $f_{k|,ij} = f_{k|,ji}$. But $f_{k|,j}\lambda_{h|}{}^{j} = 0$, $(h = 1, 2, \cdots, n;\ h \neq k)$, since the vectors $\lambda_{h|}{}^{j}$ lie in the hypersurfaces S_k. Hence, using formula (4), we conclude that

$$\cos\theta_k/b_{hk} = (\Delta_1 f_k)^{-\frac{1}{2}} f_{k|,i}(\mu_{hk|}{}^{i} - \mu_{kh|}{}^{i}). \tag{9}$$

Since any two congruences of curves C_h and C_k lie in a two-dimensional surface, the vector $\mu_{hk|}{}^{i} - \mu_{kh|}{}^{i}$ at a point P lies in the surface containing the curves C_h and C_k through P,* and hence can be written as a linear combination of the vectors $\lambda_{h|}{}^{i}$ and $\lambda_{k|}{}^{i}$.

$$\mu_{hk|}{}^{i} - \mu_{kh|}{}^{i} = a_h\lambda_{h|}{}^{i} + a_k\lambda_{k|}{}^{i}. \tag{10}$$

From (9) and (10) we conclude

THEOREM 1. *The hypersurfaces S_k are equidistant with respect to the curves C_k if and only if the vector which is the difference between the associate curvature vectors of the curves C_h and C_k with respect to one another lies always along the curves C_h for $h = 1, 2, \cdots, n;\ h \neq k$.*

In particular, if all the congruences of curves C_h, $(h = 1, 2, \cdots, n;\ h \neq k)$, are parallel with respect to the curves C_k, and if the curves C_k are parallel with respect to all the congruences of curves C_h, then the hypersurfaces S_k are equidistant with respect to the curves C_k.

From (9) also follows

THEOREM 2. *A necessary and sufficient condition that the hypersurfaces of each family be equidistant with respect to the congruences of curves not contained in them is that the associate curvature vectors of each two congruences of curves with respect to each other be identical.*

As a corollary, it follows that if the curves of each congruence are parallel with respect to the curves of every other congruence, then the hypersurfaces of each family are equidistant with respect to the congruences of curves not contained in them.†

3. *Distantial spread of one congruence of curves with respect to a second*

---

* Struik, *Grundzüge der mehrdimensionalen Differentialgeometrie*, 1922, p. 53.

† This corollary has been proved directly by Corbellini, "Di una classe di varieta caratterizzate per mezzo del parallelismo," *Rendiconti della Reale Accademia dei Lincei*, ser. 6, vol. 4 (1926), p. 94.

congruence. Let us suppose, as before, that we have n independent congruences of curves C_h, $(h = 1, 2, \cdots, n)$, with unit tangent vectors $\lambda_{h|}{}^{i}$, which are the curves of intersection of n independent families of hypersurfaces S_k, $f_k(x^1, x^2, \cdots, x^n) = c_k$, the curves C_k being taken as the transversals of the hypersurfaces S_k.

On each of the two-dimensional surfaces determined by the n-2 equations $f_i = c_i$, $(i = 1, 2, \cdots, n;\ i \neq h, k)$, there are, then, ∞^1 curves C_k, defined by the equation $f_h = c_h$, and ∞^1 curves C_h defined by the equation $f_k = c_k$. Considering the curves C_h and C_k on an arbitrary but fixed one of these surfaces, we form the logarithmic directional derivative in the positive direction of the curve C_h : $f_k = f_k{}^0$, of the distance, measured along an arbitrary curve C_k, between the curve C_h and a neighboring curve C_h : $f_k = f_k{}^0 + \Delta f_k$, and define the limit of this derivative when Δf_k approaches zero, as *the distantial spread of the curves C_h with respect to the curves C_k.*

It is readily shown that this distantial spread is given by the formula (8) and hence is identical with the distantial spread of the hypersurfaces S_k with respect to the curves C_k in the direction of the curves C_h. From (8) it follows that the curves C_h are equidistant with respect to the curves C_k in that on everyone of the two-dimensional surfaces in question each two curves C_h cut segments of equal length from the curves C_k if and only if the distantial spread of the curves C_h with respect to the curves C_k vanishes identically.

Since (9) is equivalent to (8), we have as the values of the distantial spreads of the curves C_h and C_k with respect to one another:

$$\begin{aligned} \cos\theta_k/b_{hk} &= (\Delta_1 f_k)^{-\frac{1}{2}} f_{k|,i}(\mu_{hk|}{}^{i} - \mu_{kh|}{}^{i}) \\ \cos\theta_h/b_{kh} &= (\Delta_1 f_h)^{-\frac{1}{2}} f_{h|,i}(\mu_{kh|}{}^{i} - \mu_{hk|}{}^{i}). \end{aligned} \tag{11}$$

Substituting the expression (10) in formulas (11), we find

$$a_h = -1/b_{kh}, \quad a_k = 1/b_{hk},$$

and hence

$$\mu_{hk|}{}^{i} - \mu_{kh|}{}^{i} = -(1/b_{kh})\lambda_{h|}{}^{i} + (1/b_{hk})\lambda_{k|}{}^{i}. \tag{12}$$

From this fundamental relation we obtain immediately the following theorems:

THEOREM 3. *If the curves of each of two congruences are parallel with respect to the curves of the other, then the curves of each congruence are equidistant with respect to the curves of the other.*

THEOREM 4. *If the curves of each of two congruences are equidistant*

with respect to the curves of the other, then their associate curvature vectors with respect to each other are identical.

THEOREM 5. *If the curves of each of two congruences are equidistant with respect to the curves of the other, and if the curves of one congruence of curves are parallel with respect to the other, then the curves of the second are parallel with respect to the curves of the first.*

Using Theorem 4 and remembering that the associate curvature vector of the curves C_h with respect to the curves C_k is orthogonal to the curves C_h, we have the following result:

THEOREM 6. *If at each point P the associate curvature vectors of the curves of each of two congruences with respect to the curves of the other lie in the plane determined by the tangent vectors to the curves at P, then a necessary and sufficient condition that the curves of each congruence be parallel with respect to the curves of the other is that they be equidistant with respect to the other.*

4. *Case of an orthogonal system of hypersurfaces.* Let us consider a V_n admitting an n-tuply orthogonal system of hypersurfaces S_k, $f_k = c_k$, $(k = 1, 2, \cdots, n)$. The curves of intersection C_k form an ennuple of mutually orthogonal normal congruences.

Since $\mu_{kh|}{}^i \lambda_{k|i} = 0$, and since the directions of the curves C_k coincide with the normals to the hypersurfaces S_k, we have $f_{k|,i}\mu_{kh|}{}^i = 0$. Formulas (11) then become

$$(13) \qquad \begin{aligned} 1/b_{hk} &= (\Delta_1 f_k)^{-\frac{1}{2}} f_{k|,i}\mu_{hk|}{}^i = \lambda_{k|i}\mu_{hk|}{}^i, \\ 1/b_{kh} &= (\Delta_1 f_h)^{-\frac{1}{2}} f_{h|,i}\mu_{kh|}{}^i = \lambda_{h|i}\mu_{kh|}{}^i. \end{aligned}$$

Expressing the associate curvature vectors $\mu_{hk|}{}^i$ and $\mu_{kh|}{}^i$ in terms of the coefficients of rotation γ_{pqr} of the orthogonal ennuple of curves C_k, we have

$$\mu_{hk|}{}^i = \lambda_{k|}{}^j \lambda_{h|}{}^i{}_{,j} = \lambda_{k|}{}^j \sum_{r,s} \gamma_{hrs}\lambda_{r|}{}^i \lambda_{s|j} = \sum_r \gamma_{hrk}\lambda_{r|}{}^i,$$

$$\mu_{kh|}{}^i = \lambda_{h|}{}^j \lambda_{k|}{}^i{}_{,j} = \lambda_{h|}{}^j \sum_{r,s} \gamma_{krs}\lambda_{r|}{}^i \lambda_{s|j} = \sum_r \gamma_{krh}\lambda_{r|}{}^i,$$

or

$$(14) \qquad \begin{aligned} \mu_{hk|}{}^i &= -\sum_r \gamma_{rhk}\lambda_{r|}{}^i, \\ \mu_{kh|}{}^i &= -\sum_r \gamma_{rkh}\lambda_{r|}{}^i. \end{aligned}$$

Furthermore,

$$(15) \qquad 1/r^2{}_{hk} = \sum_r \gamma^2{}_{rhk}, \quad 1/r^2{}_{kh} = \sum_r \gamma^2{}_{rkh}.$$

The formulas for the principal normal of the curves C_k,

$$\mu_{k|}{}^{i} = -\sum_{r} \gamma_{rkk}\lambda_{r|}{}^{i}, \tag{16}$$

and for the first curvature of these curves,

$$1/\rho^2{}_k = \sum_{r} \gamma^2{}_{rkk}, \tag{17}$$

are special cases of formulas (14) and (15).

From (14) we have the following result:

THEOREM 7. *A necessary and sufficient condition that the curves C_h be parallel with respect to the curves C_k is that*

$$\gamma_{rhk} = 0, \quad (r = 1, 2, \cdots, n).$$

As a special case of this theorem we have the condition that the curves C_k be geodesics: *

$$\gamma_{rkk} = 0, \quad (r = 1, 2, \cdots, n).$$

Formulas (14) through (17) and Theorem 7 hold for any orthogonal ennuple of congruences whatsoever. In this particular case, since the congruences of curves are all normal, the coefficients of rotation γ_{rhk} vanish for values of r, h, and k all distinct. Hence formulas (14) become

$$\mu_{hk|}{}^{i} = \gamma_{hkk}\lambda_{k|}{}^{i}, \quad \mu_{kh|}{}^{i} = \gamma_{khh}\lambda_{h|}{}^{i}. \tag{18}$$

Putting these values in equations (13) we have

$$1/b_{hk} = \gamma_{hkk}, \qquad 1/b_{kh} = \gamma_{khh}. \tag{19}$$

Since γ_{khk} (which is the negative of γ_{hkk}) is the orthogonal projection on C_h of the first curvature vector of C_k, we have

THEOREM 8. *The distantial spread at a point P of the curves C_h of a congruence with respect to the curves C_k of another is the negative of the orthogonal projection of the first curvature vector at P of the curve C_k on the tangent at P to the curve C_h.*

Formulas (18) can now be rewritten as

$$\mu_{hk|}{}^{i} = (1/b_{hk})\lambda_{k|}{}^{i}, \quad \mu_{kh|}{}^{i} = (1/b_{kh})\lambda_{h|}{}^{i}. \tag{20}$$

Hence we conclude the following theorems:

THEOREM 9. *The absolute value of the distantial spread of the curves of one congruence with respect to the curves of a second is equal to the associate*

* Eisenhart, *Riemannian Geometry*, p. 100.

curvature of the curves of the first congruence with respect to those of the second.

THEOREM 10. *If the curves of one congruence are equidistant with respect to the curves of another, then the curves of the first congruence are also parallel with respect to the curves of the second, and conversely.*

THEOREM 11. *The associate directions of the curves of one congruence with respect to the curves of a second coincide with the direction of the latter curves.*

From formula (18) it follows that

$$1/r^2_{hk} = \gamma^2_{hkk}, \qquad (h \neq k).$$

Summing over h, omitting the value $h = k$, we have by formula (17)

$$\sum_h 1/r^2_{hk} = \sum_h \gamma^2_{hkk} = 1/\rho^2_k.$$

Hence

$$\sum_h 1/b^2_{hk} = \sum_h 1/r^2_{hk} = 1/\rho^2_k. \qquad (h \neq k). \tag{21}$$

Considering this result from the point of view of the distantial spread of the family of hypersurfaces S_k with respect to their orthogonal trajectories C_k, we note that three properties are involved here:

(a) The hypersurfaces S_k equidistant with respect to the curves C_k.

(b) The curves C_k geodesic.

(c) The curves C_h, $(h = 1, 2, \cdots, n;\ h \neq k)$, parallel with respect to the curves C_k.

From (21) we have

THEOREM 12. *If one of these three properties is valid, then so are the other two.**

5. *Oblique system of hypersurfaces.* We return to the general case when the n families of hypersurfaces are not orthogonal, and introduce the notation $1/\rho_{hk}$ for the projection of the associate curvature vector $\mu_{hk|}{}^i$ on $\lambda_{k|}{}^i$. Multiplying equation (12) first by $g_{ij}\lambda_{k|}{}^j$, and then by $g_{ij}\lambda_{h|}{}^j$, summing over i and j, we obtain the relations

$$\begin{aligned} 1/p_{hk} &= 1/b_{hk} - \cos\Omega_{hk}/b_{kh}, \\ 1/p_{kh} &= -\cos\Omega_{hk}/b_{hk} + 1/b_{kh}, \end{aligned} \tag{22}$$

* This result for the case when property (b) is given is not new. See Eisenhart, *loc. cit.*, p. 57. Such families of hypersurfaces are said to be geodesically parallel.

where Ω_{hk} is the angle between the curves C_h and C_k. Solving these two equations for $1/b_{hk}$ and $1/b_{kh}$, we find

$$(23)\qquad \begin{aligned} \sin^2\Omega_{hk}/b_{hk} &= 1/p_{hk} + \cos\Omega_{hk}/p_{kh}, \\ \sin^2\Omega_{hk}/b_{kh} &= \cos\Omega_{hk}/p_{hk} + 1/p_{kh}. \end{aligned}$$

These results show that we have to deal with the following four properties of two congruences of curves:

$$(A)\qquad \begin{cases} \text{Curves } C_h \text{ equidistant with respect to the curves } C_k\text{: } 1/b_{hk} = 0. \\ \text{Associate direction of the curves } C_h \text{ with respect to the curves } C_k \\ \qquad \text{orthogonal to the curves } C_k \text{ : } 1/p_{hk} = 0. \end{cases}$$

$$\begin{cases} \text{Curves } C_k \text{ equidistant with respect to the curves } C_h \text{ : } 1/b_{kh} = 0. \\ \text{Associate direction of the curves } C_k \text{ with respect to the curves } C_h \\ \qquad \text{orthogonal to the curves } C_h \text{ : } 1/p_{kh} = 0. \end{cases}$$

From equations (21) and (22) we conclude:

THEOREM 13. *Two non-orthogonal congruences of curves which have any two of the properties* (A) *have the other two also.*

The curves C_h and C_k may form an orthogonal system of curves without the families of hypersurfaces being an n-tuply orthogonal system. For such curves we have:

THEOREM 14. *Two orthogonal congruences of curves having one of the first two properties* (A) *and one of the last two have all four properties.*

Finally we have

THEOREM 15. *If the distantial spreads of the curves of each congruence with respect to the curves of the other are equals or negatives of each other, then so are the projections on the curves of each congruence of the associate curvature vectors, with respect to those curves, of the curves of the other congruence, and conversely.*

Using the quantities $1/p_{hk}$ and $1/p_{kh}$ we can express the components of the associate directions as follows:

$$(24)\qquad \begin{aligned} \sin^2\Omega_{hk}\mu_{hk|}{}^{j} &= -(\cos\Omega_{hk}/p_{hk})\lambda_{h|}{}^{j} + (1/p_{hk})\lambda_{k|}{}^{j} + \nu_{hk|}{}^{j}, \\ \sin^2\Omega_{hk}\mu_{kh|}{}^{j} &= (1/p_{kh})\lambda_{h|}{}^{j} - (\cos\Omega_{hk}/p_{kh})\lambda_{k|}{}^{j} + \nu_{kh|}{}^{j}, \end{aligned}$$

where $\nu_{hk|}{}^{j} = \nu_{kh|}{}^{j}$ is the common component of the associate directions orthogonal to the plane of $\lambda_{h|}{}^{j}$ and $\lambda_{k|}{}^{j}$.

RADCLIFFE COLLEGE.

ON A RATIONAL SURFACE OF ORDER 12 IN 9-SPACE AND ITS PROJECTIONS.

By B. C. Wong.

1. *Introduction.* The quadratic transformation *

$$\rho x_{ij} = x_i x_j \qquad [i, j = 1, 2, 3, 4] \tag{1}$$

where x_1, x_2, x_3, x_4 are the homogeneous coördinates of a point of a 3-space S_3 and $x_{11}, \cdots, x_{34}$ are those of a 9-space S_9 transforms the points of S_3 into the points of an octavic variety V_3^8 of order 8 and dimension 3 in S_9. To a plane of S_3 corresponds a Veronese quartic surface † lying on V_3^8 and to a quadric surface of S_3 corresponds a section of V_3^8 by a hyperplane of S_9. Of interest is the surface Φ^{12} of order 12 on V_3^8 which corresponds to a general cubic surface F^3 of S_3. It has 27 conics whose relative positions can best be studied by reference to the relative positions of the 27 lines on F^3. By representing F^3 upon a plane ϕ by means of the ∞^3 cubic curves through six general points $P_1, P_2, \cdots, P_6$ of ϕ, we see that Φ^{12} can be represented upon ϕ by means of the ∞^9 sextic curves with six nodes at P_λ $[\lambda = 1, 2, \cdots, 6]$.

If we let $(y_1 : y_2 : y_3)$ be the coördinates of a point of ϕ and $(p_1^{(\lambda)} : p_2^{(\lambda)} : p_3^{(\lambda)})$ be those of the six points P_λ, we may take the equations

$$f_i(y_1 : y_2 : y_3) \equiv \sum_1^3 u_{klm} y_k y_l y_m = 0 \qquad [i = 1, 2, 3, 4]$$

where

$$f_i(p_1^{(\lambda)} : p_2^{(\lambda)} : p_3^{(\lambda)}) \equiv 0$$

to be the equations of four general cubics of the ∞^3-system $| c^3 |$ of cubics passing through the six points P_λ in ϕ. Then the equations

$$x_1 : x_2 : x_3 : x_4 = f_1 : f_2 : f_3 : f_4 \tag{2}$$

are those of the cubic transformation which transforms a point of ϕ into a point of F^3. They are the parametric equations of F^3.

* This transformation was mentioned in B. C. Wong, "A hypersurface of order 2^{r-1} in r-space," *American Journal of Mathematics*, vol. 54 (1932), pp. 293-298. It was used by W. G. Welchman in his paper, "On elliptic quartic curves with assigned points and chords," and by J. A. Todd in his paper, "Some enumerative results for elliptic quartic curves," *Proceedings of the Cambridge Philosophical Society*, vol. 27 (1931), pp. 20-23 and 538-542 respectively, to solve certain problems on elliptic quartic curves.

† Bertini, *Projective Geometrie Mehrdimensionaler Räume* (1924), Chapter 16.

Now the equation of a general sextic curve of the ∞^9-system $|c^6|$ of sextic curves having the points P_λ for nodes is of the form

$$\sum_1^4 A_{ij} f_i f_j = 0.$$

This sextic c^6 has for image in S_3 the sextic curve C^6 in which F^3 is met by the quadric surface

$$\sum_1^4 A_{ij} x_i x_j = 0.$$

Now corresponding to C^6 is the curve Γ^{12} of order 12 which is the intersection of Φ^{12} and the hyperplane

$$\sum_1^4 A_{ij} x_{ij} = 0.$$

Then Γ^{12} is the image of c^6 in ϕ under the correspondence, whose equations are

$$\sigma x_{ij} = f_i f_j.$$

These are also the parametric equations of Φ^{12}.

If we now project ϕ^{12} from a general 5-space of S_9 upon a 3-space, in particular the 3-space S_3 containing F^3, we have for projection a surface Φ'^{12} which is represented upon ϕ by the sextic curves of a certain ∞^3-system $|c^6|'$ of $|c^6|$. If the center of projection is in a particular position relative to Φ^{12}, the projection will possess particular properties or will be of lower order, say $12 - \nu$. The properties of these different projections can be described without difficulty. The question arises as to how to choose the center of projection so that the curves of $|c^6|'$ will be the Jacobians of the nets of cubics of $|c^3|$. If the selection is properly made, the surface Φ'^{12} will be reciprocal to F^3. The sextic curves C^6, which are the images of the curves of $|c^6|'$ under the transformation (2) and which have for images under the transformation (1) the sections of Φ^{12} by the hyperplanes of S_9 through the center of projection, will be the intersections of F^3 and the polar quadric surfaces of all the points of S_3 with respect to F^3.

In what follows we shall, after describing very briefly some of the various projections of Φ^{12}, determine the 5-space S_5 for center of projection so that the projection Φ'^{12} will be reciprocal to F^3. Then, as a special case, we shall obtain the Steiner's quartic surface from the Φ^{12} corresponding to the four-nodal cubic surface in S_3. Finally, we shall obtain from the same Φ^{12} a surface reciprocal to a certain nonic surface whose equation is

$$x_1^{1/3} + x_2^{1/3} + x_3^{1/3} + x_4^{1/3} = 0.$$

2. *Properties of* $\Phi'^{12-\nu}$. If the 5-space S_5, the center of projection, has ν $[0 \leqq \nu \leqq 6]$ general points in common with the Φ^{12} corresponding to a general cubic surface F^3 of S_3, the projection of Φ^{12} is a surface $\Phi'^{12-\nu}$ of order $12-\nu$. It has a double curve of order $b=(\nu^2-21\nu+102)/2$, $j=42-4\nu$ pinch points and $t=(10-\nu)(9-\nu)(8-\nu)/6+3(\nu-7)$ triple points. It is of class $m=33$ for all permissible values of ν. If $\nu=0$, then, for Φ'^{12}, $b=51$, $j=42$, and $t=99$. Now let S_5 be incident with μ planes, each containing a conic of Φ^{12}, and keeping $\nu=0$ we see that the double curve is composed of a curve of order $51-\mu$ and μ lines on each of which are two of the 42 pinch points.

Now let S_5 contain a conic of Φ^{12} and have ν' $[0 \leqq \nu' \leqq 3]$ general points in common with Φ^{12}, and we have for projection a surface $\Phi'^{7-\nu'}$ of order $7-\nu'$ and class 20 on which lies a double curve of order $b=(\nu'^2-11\nu'+26)/2$. There are $j=4(4-\nu')$ pinch points and $t=(2-\nu')(3-\nu')(7-\nu')/6$ triple points. For $\nu'=3$, we have the familiar rational quartic surface with a double line.* If S_5 contains two non-intersecting conics of Φ^{12}, then the projection is a quadric surface; but if S_5 contains two intersecting conics, the projection is a quartic surface with a double conic.† Finally, if S_5 contains an elliptic sextic curve of Φ^{12} or contains three conics of Φ^{12} which are incident in pairs, the projection is a general cubic surface.

The results above can be verified easily by considering the representation of each of the projections upon the plane ϕ or by means of known formulas concerning rational surfaces.‡

3. *The determination of* S_5 *so that* Φ'^{12} *will be reciprocal to* F^3. Denote the 27 conics on Φ^{12} by the symbols $\alpha_1, \cdots, \alpha_6, \beta_1, \cdots, \beta_6, \gamma_{12}, \gamma_{13}, \cdots, \gamma_{56}$ in conformity with the double-six notation $a_1, \cdots, a_6$, $b_1, \cdots, b_6$, $c_{12}, c_{13}, \cdots, c_{56}$ for the 27 lines on F^3. The 27 conics are such that each intersects 10 others each in one point and that they all form 45 triples, the members of each triple being incident in pairs. The planes of non-incident conics are skew.

Consider, in particular, the conic γ_{ij}. It intersects γ_{kl} in the point $\gamma_{ij}\cdot\gamma_{kl}$ and γ_{mn} in the point $\gamma_{ij}\cdot\gamma_{mn}$ $[i \neq j \neq k \neq l \neq m \neq n]$ while γ_{kl} intersects γ_{mn} in the point $\gamma_{kl}\cdot\gamma_{mn}$. Denote the plane of these three points by $(\gamma_{ij}\cdot\gamma_{kl}\cdot\gamma_{mn})$. By permuting the subscripts we obtain 15 such planes. Now γ_{ij} intersects the conic α_i in a point $\gamma_{ij}\cdot\alpha_i$ and the conic β_j in a point

* Jessop, *Quartic Surfaces*, Cambridge Press, 1916, Chapter VI.

† *Ibid.*, Chapter IV.

‡ Cayley, *Collected Mathematical Papers*, vol. 8, pp. 388-393.

$\gamma_{ij} \cdot \beta_j$ while α_i and β_j intersect in a point $\alpha_i \cdot \beta_j$. Denote the planes of these three points by $(\gamma_{ij} \cdot \alpha_i \cdot \beta_j)$. Thus, we get 30 planes of this type.

Now the points on each of the 27 conics corresponding to the points on each of the 27 lines of F^3, where the planes through the line are doubly tangent to F^3, are paired in involution. To the involution on γ_{ij} belong the three pairs of points $\gamma_{ij} \cdot \gamma_{kl}$, $\gamma_{ij} \cdot \gamma_{mn}$; $\gamma_{ij} \cdot \gamma_{km}$, $\gamma_{ij} \cdot \gamma_{ln}$; $\gamma_{ij} \cdot \gamma_{kn}$, $\gamma_{ij} \cdot \gamma_{lm}$ and also the two pairs $\gamma_{ij} \cdot \alpha_i$, $\gamma_{ij} \cdot \beta_j$; $\gamma_{ij} \cdot \alpha_j$, $\gamma_{ij} \cdot \beta_i$. The lines determined by all these pairs are all concurrent at a point C_{ij}, the center of involution. In a similar manner we obtain the centers C_{kl} and C_{mn} of involution for the conics γ_{kl} and γ_{mn} respectively. These three points C_{ij}, C_{kl}, C_{mn} lie on a line which may be denoted by $C_{ij} \cdot C_{kl} \cdot C_{mn}$ and this line lies in the plane $(\gamma_{ij} \cdot \gamma_{kl} \cdot \gamma_{mn})$. There are 15 such lines. Let the centers of involution on the conics α_i, β_j be denoted by A_i, B_j respectively. The three points C_{ij}, A_i, B_j lie on a line $C_{ij} \cdot A_i \cdot B_j$ of the plane $(\gamma_{ij} \cdot \alpha_i \cdot \beta_j)$. There are 30 such lines.

The 27 points A_i, B_j, C_{ij} and the 45 lines just obtained form a configuration such that each point is on five lines and each line contains three points. Now we wish to show that this configuration lies completely in a 5-space. Consider two skew lines, say $C_{12} \cdot C_{34} \cdot C_{56}$ and $C_{13} \cdot C_{25} \cdot C_{46}$, of the configuration. They determine four other lines, namely, $C_{12} \cdot C_{46} \cdot C_{35}$, $C_{13} \cdot C_{56} \cdot C_{24}$, $C_{34} \cdot C_{25} \cdot C_{16}$, $C_{24} \cdot C_{35} \cdot C_{16}$. All these six lines lie in a 3-space S_3. Now choose a line through one of the nine points already in S_3, say the line $C_{12} \cdot C_{26} \cdot C_{45}$ which meets S_3 in the point C_{12}. This line determines with S_3 a 4-space S_4. Now it is easy to see that this S_4 contains all the 15 lines of the type $C_{ij} \cdot C_{kl} \cdot C_{mn}$. But the line $C_{12} \cdot A_1 \cdot B_2$, having the point C_{12} in common with S_4, determines with it a 5-space S_5. In this S_5 are obviously the lines $C_{13} \cdot A_1 \cdot B_3$, $C_{13} \cdot A_3 \cdot B_1$, $C_{14} \cdot A_1 \cdot B_4$, and the rest of the 30 lines of the type $C_{ij} \cdot A_i \cdot B_j$.

This is the very S_5 we have been seeking. Taking it as the center of projection, we see that the 27 double lines on the projection Φ'^{12} come from the 27 conics of Φ^{12} and that the 45 triple points come from the 45 triples of points on Φ^{12} of which 15 are of the type $\gamma_{ij} \cdot \gamma_{kl}$, $\gamma_{ij} \cdot \gamma_{mn}$, $\gamma_{kl} \cdot \gamma_{mn}$ and 30 are of the type $\gamma_{ij} \cdot \alpha_i$, $\gamma_{ij} \cdot \beta_j$, $\alpha_i \cdot \beta_j$. There is no need of describing the properties of Φ'^{12} as they are already familiar.*

4. *The Steiner's quartic surface.* As a very special case we consider the Φ^{12} in S_9 corresponding to a four-nodal cubic surface F^3. Let F^3 have the equation

---

* Salmon-Rogers, *Analytic Geometry of Three Dimensions*, 5th ed., vol. 2, Chapter XVIIa.

$$1/x_1 + 1/x_2 + 1/x_3 + 1/x_4 = 0$$

or the equations

$$x_1 : x_2 : x_3 : x_4 = 1/y_1 : 1/y_2 : 1/y_3 : 1/y_4$$

where

$$y_1 + y_2 + y_3 + y_4 \equiv 0.$$

The base points of the cubic curves of representation in ϕ are the six vertices of the quadrilateral y_1, y_2, y_3, y_4. Regarding

$$(3) \qquad \begin{aligned} f_1 &\equiv y_2y_3y_4 = 0, \\ f_2 &\equiv y_3y_4y_1 = 0, \\ f_3 &\equiv y_4y_1y_2 = 0, \\ f_4 &\equiv y_1y_2y_3 = 0 \end{aligned}$$

as four independent cubics of the ∞^3-system $|\, c^3 \,|$, we have

$$(4) \qquad \begin{aligned} \Sigma A_{ij}f_if_j \equiv\ & A_{11}y_2^2y_3^2y_4^2 + A_{22}y_3^2y_4^2y_1^2 + A_{33}y_4^2y_1^2y_2^2 + A_{44}y_1^2y_2^2y_3^2 \\ & + 2y_1y_2y_3y_4(A_{12}y_3y_4 + A_{13}y_4y_2 + A_{23}y_1y_4 \\ & + A_{34}y_1y_2 + A_{24}y_1y_3 + A_{14}y_2y_3) = 0 \end{aligned}$$

for the equation of the ∞^9-system $|\, c^6 \,|$ of sextics having nodes at the base points of $|\, c^3 \,|$. The equation of Φ^{12} may be written

$$\sigma x_{ij} = 1/y_iy_j.$$

Now it is easy to see that Φ^{12} has four conical points given by

$$x_{ii} = 1, \quad x_{ij} = x_{kj} = x_{kl} = 0 \qquad [i, j = 1, 2, 3, 4].$$

Projecting Φ^{12} from any 5-space containing these four conical points upon S_3, we obtain a Steiner's quartic surface. But we wish to project from a particular 5-space so that the projection will be the exact reciprocal of F^3, that is, will have the equation

$$x_1^{\frac{1}{2}} + x_2^{\frac{1}{2}} + x_3^{\frac{1}{2}} + x_4^{\frac{1}{2}} = 0.$$

There are nine conics on Φ^{12} of which six intersect three by three on the conical points and the remaining three are incident in pairs. Denote the first six conics by $\beta_{12}, \beta_{13}, \cdots, \beta_{34}$ where β_{ij} passes through the conical points

$$X_{ii}(x_{ii} = 1,\ x_{ij} = x_{kj} = x_{kl} = 0) \quad \text{and} \quad X_{jj}(x_{jj} = 1,\ x_{ij} = x_{kj} = x_{kl} = 0)$$

and the other three by $\gamma_{12,34}$, $\gamma_{13,42}$, $\gamma_{14,23}$ where $\gamma_{ij,kl}$ is incident with β_{ij}

and β_{kl}. The points on each of these three conics, corresponding to the points on each of the three coplanar lines of F^3 not passing through the nodes where the planes through the line are doubly tangent to F^3, are paired in involution. Let $C_{12,34}$, $C_{13,42}$, $C_{14,23}$ be the centers of involutions on the conics $\gamma_{12,34}$, $\gamma_{13,42}$, $\gamma_{14,23}$ respectively. They have the coördinates, as a little computation will show,

$$x_{11} : x_{22} : x_{33} : x_{44} : x_{12} : x_{34} : x_{13} : x_{42} : x_{14} : x_{23}$$
$$C_{12,34}\ (0 : 0 : 0 : 0 : 0 : 0 : 1 : 1 : -1 : -1),$$
$$C_{13,24}\ (0 : 0 : 0 : 0 : 1 : 1 : 0 : 0 : -1 : -1),$$
$$C_{14,23}\ (0 : 0 : 0 : 0 : 1 : 1 : -1 : -1 : 0 : 0).$$

These three points evidently lie on a line. Now the 5-space containing the four conical points and this line is the S_5 we desire to have for center of projection. Every hyperplane through this S_5 has an equation of the form

$$(5) \qquad A_{12}x_{12} + A_{34}x_{34} + A_{13}x_{13} + A_{42}x_{42} + A_{14}x_{14} + A_{23}x_{23} = 0$$

where

$$A_{12} + A_{34} = A_{13} + A_{42} = A_{14} + A_{23}.$$

Now let

$$A_{12} : A_{34} : A_{13} : A_{42} : A_{14} : A_{23}$$
$$= u_3 + u_4 : u_1 + u_2 : u_4 + u_2 : u_1 + u_3 : u_2 + u_3 : u_1 + u_4$$

and equation (5) becomes

$$(u_3 + u_4)x_{12} + (u_1 + u_2)x_{34} + (u_4 + u_2)x_{13} + (u_1 + u_3)x_{42} + (u_2 + u_3)x_{14} + (u_1 + u_4)x_{23} = 0.$$

Putting these values of the coefficients and also $A_{ii} = 0$ in equation (4), we have, after reduction,

$$u_1y_1^2 + u_2y_2^2 + u_3y_3^2 + u_4y_4^2 = 0$$

for the equation of the Jacobians of the nets of the ∞^3-system $|\, c^3 \,|$ of cubic curves representing the four-nodal cubic surface F^3. These Jacobian conics represent the plane sections of the Steiner's quartic surface whose equation is evidently

$$x_1^{1/2} + x_2^{1/2} + x_3^{1/2} + x_4^{1/2} = 0.$$

5. *The Φ'^{12} reciprocal to a certain nonic surface.* The totality of cubic curves in the plane ϕ may be taken to represent the normal rational surface F^9 in S_9. Projecting F^9 from a certain S_5 upon S_3, we obtain a surface F'^9

which may be represented upon ϕ by the system of cubic curves given by the equation

$$A_1 y_1^3 + A_2 y_2^3 + A_3 y_3^3 + A_4 y_4^3 = 0,$$

where

$$y_1 + y_2 + y_3 + y_4 \equiv 0,$$

and therefore has the equation

$$x_1^{1/3} + x_2^{1/3} + x_3^{1/3} + x_4^{1/3} = 0.$$

Consider again the Φ^{12} corresponding to the four-nodal cubic surface F^3. In order to obtain a Steiner's quartic surface for projection we took for center of projection an S_5 containing the four conical points of Φ^{12}. Now suppose we take for center of projection the S_5 determined by the six vertices other than X_{ii} of the coördinate simplex given by the coördinates

$$x_{ii} = 0, \quad x_{ij} = 1, \quad x_{ik} = x_{jk} = x_{kl} = 0.$$

The ∞^3 hyperplanes through this S_5 have for equation

$$A_{11}x_{11} + A_{22}x_{22} + A_{33}x_{33} + A_{44}x_{44} = 0$$

and the Jacobian sextic curves of the nets of the ∞^3-system of cubic curves given by (6) have for equation

$$A_{11}/y_1^2 + A_{22}/y_2^2 + A_{33}/y_3^2 + A_{44}/y_4^2 = 0.$$

These sextics are the curves of representation upon ϕ of the projection Φ'^{12} of Φ^{12} reciprocal to F'^9. The equation of Φ'^{12} is evidently

$$x_1^{-1/2} + x_2^{-1/2} + x_3^{-1/2} + x_4^{-1/2} = 0.$$

The properties of F'^9 have been studied recently by A. R. Williams * and those of Φ'^{12} can be easily deduced either by dualizing F'^9 or from its representation upon ϕ.

UNIVERSITY OF CALIFORNIA.

* "Analogues of the Steiner surface and their double curves," *Bulletin of the American Mathematical Society*, vol. 39 (1933), pp. 621-626.

REDUCIBLE EXCEPTIONAL CURVES OF THE FIRST KIND.*

By S. F. Barber † and Oscar Zariski

Introduction. In a birational transformation between two surfaces to certain curves of one of the surfaces may correspond *simple* points of the other. Such curves have been called *exceptional curves.* Exceptional curves have been subdivided into those of the first or second kind according as a point of the curve is not or is transformed into a curve of the other surface. The ones previously considered have either been irreducible or consisted of at most two components. A treatment of reducible exceptional curves with components all simple is found in ([4], Chap. II, 6). We shall consider exceptional curves of the first kind with s irreducible components, each counted a certain number of times. This case arises when the birational transformation possesses infinitely near fundamental points on one or several irreducible algebroid branches.

In the sequel we shall use some fundamental notions of linear systems of curves on an algebraic surface and the theory of singularities as developed by Enriques ([3], pp. 327-399).

Sections 1 and 2 of this paper are devoted to definitions and the derivation of the virtual degree and genus of an exceptional curve of the first kind; the presentation here is parallel to but more complete than that found in ([4], Chap. II, 6). In sections 3 and 4 we introduce s fundamental points of the birational transformation, mentioned above, on that surface where the given exceptional curve has been changed into a point, and limit ourselves to the case in which these fundamental points lie on a single irreducible algebroid branch. In 5 we consider the correspondence between the irreducible components of the given curve on the one surface and the immediate neighborhood of each of the fundamental points on the other surface. In 6 we determine the multiplicities of these fundamental points on branches of lowest order passing through the first α ($\alpha \leq s$) of them. In 7 we describe the intersections of the irreducible components of the given exceptional curve, using extensively the notion of proximate points introduced by Enriques in his theory of singularities ([3], p. 381); in 8 we study the classification of the s fundamental points as free points or satellites when the intersection numbers

* Presented to the Society, March 30, 1934.

† National Research Fellow.

of the components of the exceptional curve are known. In 9 we extend to reducible exceptional curves of the first kind the characterization of irreducible exceptional curves of the first kind as found in (1). Finally in 10 we consider the situation when the fundamental points lie on several irreducible algebroid branches.

1. *Definitions.* Let F be an algebraic surface in S_ρ, free from singularities, and let Σ be a linear system of curves C on F, of dimension r. We shall say for brevity that a curve L on F, irreducible or not, is a total fundamental curve of Σ, if L is a fixed component of a linear subsystem, ∞^{r-1}, of Σ and if the curves of this subsystem do not have other fixed components outside L. We shall adopt the following

Definition. (4, Chap. II, 6). *A curve L on F is an exceptional curve if there exists on F a linear system Σ of curves C such that*

(i) *Σ is irreducible, simple, of dimension $r \geq 3$;*

(ii) *L is a total fundamental curve of Σ;*

(iii) *The ∞^{r-1} curves C_1, which together with L give total curves of Σ, form a linear system Σ_1 of effective degree one less than the effective degree of Σ.*

We recall that the effective degree of a linear system is the number of *variable* intersections of two curves of the system. If the system is of dimension r, its effective degree is $\geq r-1$. It follows that the system Σ_1 is necessarily irreducible, since by (ii) it is free from fixed components and by (i) ($r \geq 3$) and (iii) it is of positive effective degree and therefore cannot be composed of the curves of a pencil.

If it is possible to find the above system Σ in such a manner that it should have no base points on L, then L is called an exceptional curve of the *first kind.* If this is not possible, L is called an exceptional curve of the *second kind.*

If we refer the ∞^r curves C of Σ to the hyperplanes of an S_r, we obtain in S_r a surface $\bar{F}$ birationally equivalent to F, since Σ is simple. This surface is of order n, where n is the effective degree of Σ, and its hyperplane sections $\bar{C}$ correspond to the curves C of Σ. In this correspondence to the subsystem $\Sigma_1 + L$ of Σ there corresponds the system of hyperplane sections of $\bar{F}$ on a fixed point O of S_r. This point O is on $\bar{F}$, since the effective degree of Σ_1 is less than the effective degree of Σ, and is a *simple* point of $\bar{F}$, since by (iii) two hyperplane sections of $\bar{F}$ on O have $n-1$ variable intersections. Thus in the birational correspondence T between F and $\bar{F}$ the curve L (irreducible or not) is transformed into a *simple* point O of $\bar{F}$ (a fundamental point of T^{-1}) and this is in agreement with the usual definition of an exceptional curve.

It is one of the purposes of this paper to study the nature of the correspondence between L and the neighborhood of O when L is reducible.

The fundamental points of T (on F) are at the base points of Σ. Hence an exceptional curve L is of the first or of the second kind (see above definition) according as it is or is not possible to transform L into a simple point of a surface $\bar{F}$, birationally equivalent to F, in such a manner that no point of L is at the same time transformed into a curve of $\bar{F}$.

In the sequel we shall be concerned only with exceptional curves of the first kind. We observe, however, that the properties of exceptional curves of the second kind can be deduced from those of the first kind. In fact, it is always possible to transform birationally F into a surface $\bar{F}$ in such a manner that the transformation should possess no fundamental points on $\bar{F}$ and that to Σ there should correspond on $\bar{F}$ a linear system $\bar{\Sigma}$, free from base points. For this it is sufficient to refer the sections of F by hyperquadrics on a base point of Σ to the hyperplanes of a linear space and to apply this procedure successively to the transformed surfaces until all the base points of Σ disappear. To the exceptional curve L there will correspond on $\bar{F}$ an exceptional curve $\bar{L}$ of the *first kind*. However, in general $\bar{L}$ will consist of the transform proper of L and of the fundamental curves of the transformation on $\bar{F}$, arising from the base points of Σ.

2. *The virtual characters of an exceptional curve of the first kind.* *Theorem* 1 ([4], Chap. II, 6). *If L is an exceptional curve of the first kind, then $(L^2) = -1$ and $[L] = 0$.**

Proof. Let H denote the set of effective base points of Σ. We have $(C \cdot L) = ((C_1 + L) \cdot L) = 0$, since L, a total fundamental curve of Σ, does not have variable intersections with a C and since no base point of Σ is on L. Consequently $(C_1 \cdot L) = -(L^2)$. Now $C_1 + L$ is the limit of an irreducible curve C and hence C_1 and L must have at least one point in common (Principle of degeneration of Enriques[2]). It follows that $(C_1 \cdot L) > 0$ and hence $(L^2) < 0$. We also have† $n = (C^2)_H = (C_1^2)_H + 2(C_1 \cdot L) + (L^2) = (C_1^2)_H - (L^2)$. From the hypothesis that the effective degree of Σ_1 is $n-1$, it follows that $(C_1^2)_H \geq n-1$,‡ and that consequently $(L^2) \geq -1$.

---

* (L^2) is the virtual degree, $[L]$ the virtual genus. In evaluating these characters of L, we consider L as *virtually free from base points.*

† $(C^2)_H$ denotes the virtual degree of C with respect to the assigned set H of base points, in this case the effective degree of Σ.

‡ The set of effective base points of Σ_1 certainly includes H, since $C_1 + L = C$. *A priori* we cannot exclude the possibility that Σ_1 may possess accidental base points, whence the necessity of using the inequality sign.

It follows that $(L^2) = -1$ and incidentally

$$(C_1^2)_H = n - 1; \tag{1}$$

$$(C_1 \cdot L) = 1. \tag{2}$$

We observe that from (1) it follows that H is also the set of effective base points of Σ_1 and hence $[C_1]_H$ gives the effective genus of C_1. Since the curves C_1 correspond to the hyperplane sections of $\bar{F}$ on a simple point O, it follows that $[C_1]_H = [C]_H$. Furthermore

$$[C]_H = [C_1]_H + [L] + (C_1 \cdot L) - 1 = [C_1]_H + [L].$$

Hence $[L] = 0$, q. e. d.

The exceptional curve L may be reducible and some of its irreducible components may occur in L to a certain multiplicity. Let then $L = k_1L_1 + k_2L_2 + \cdots + k_sL_s$, where $L_1, L_2, \cdots, L_s$ are the distinct irreducible components of L and where $k_1, k_2, \cdots, k_s$ are positive integers. From (2) it follows that C_1 intersects one and only one of the components L_i and that this must be a simple component of L. Choosing our notation properly we may assume that

$$(C_1 \cdot L_1) = 1, \quad (C_1 \cdot L_i) = 0 \qquad (i = 2, \cdots, s), \tag{3}$$

$$k_1 = 1. \tag{3'}$$

Since $(C \cdot L_i) = 0$ $(i = 1, \cdots, s)$, it follows that

$$(L \cdot L_1) = -1, \quad (L \cdot L_i) = 0 \qquad (i = 2, \cdots, s), \tag{4}$$

all curves L_i being considered as virtually free from base points. The relations (4) can be rewritten as follows:

$$\begin{aligned} &(L_1^2) + \sum_{i=2}^{s} k_i (L_1 \cdot L_i) = -1, \\ &(L_j \cdot L_1) + \sum_{i=2}^{s} k_i (L_j \cdot L_i) = 0 \qquad (j = 2, \cdots, s). \end{aligned} \tag{5}$$

We observe that the curve $L_1 + L_2 + \cdots + L_s$ must be *connected*, i. e. it cannot be represented as the sum of two curves having no points in common. This follows immediately from the fact that C_1 meets only the curve L_1 and that $C_1 + L$ is the limit of the irreducible curve C. Since $(L_i \cdot L_j) \geqq 0$, if $i \neq j$, and since, by the preceding remark, for a given value of i one at least of the intersection numbers $(L_i \cdot L_j)$, $i \neq j$, must be positive, it follows from (5) *that the virtual degree* (L_i^2) *of any component* L_i *of* L *is negative.*

Applying the well known formula for the virtual genus of a reducible curve, we find

$$[L] = \sum_{i=1}^{s} k_i p_i + \sum_{i=1}^{s} [k_i(k_i - 1)/2](L_i)^2 + (1/2) \sum_{\substack{i=1 \\ i \neq j}}^{s} \sum_{j=1}^{s} k_i k_j (L_i \cdot L_j) - \sum_{i=1}^{s} k_i + 1,$$

where $p_i = [L_i]$. By means of (4) or the equivalent relations (5) this expression of $[L]$ can be simplified:

$$[L] = \sum_{i=1}^{s} k_i p_i - k_1/2 - (1/2) \sum_{i=1}^{s} k_i (L_i^2) - \sum_{i=1}^{s} k_i + 1.$$

Since by Theorem 1 we have $[L] = 0$ and since $k_1 = 1$ we arrive at the following relation:

$$\sum_{i=1}^{s} k_i [2p_i - (L_i^2) - 2] + 1 = 0. \tag{6}$$

3. *The fundamental points in the first neighborhood of O.* We now proceed to study the system Σ_1, the residual of Σ with respect to L. By (3) the curves $L_2, \cdots, L_s$ are fundamental curves of Σ_1. On the contrary L_1 is not a fundamental curve of Σ_1, since $(C_1 \cdot L_1) = 1$, and since the intersection of C_1 and L_1 cannot be a fixed intersection, because by (1) the effective base points of Σ_1 coincide with those of Σ and are all outside L. Let us assume for the sake of generality that $L_2 + \cdots + L_s$ consists of several connected curves $\bar{L}^{(1)}, \bar{L}^{(2)}, \cdots \bar{L}^{(\sigma)}$, having two by two no points in common (maximal connected components of $L_2 + \cdots + L_s$). Each $\bar{L}^{(i)}$ is then itself a fundamental curve of Σ_1. Let us consider for instance $\bar{L}^{(1)}$ and let us denote by Σ_2 the residual system of Σ_1 with respect to $\bar{L}^{(1)}$. Let C_2 denote a generic curve of Σ_2. What are the possible fixed components of Σ_2? L_1 is not a fixed component since it is not fundamental for Σ_1. Neither can any component of $\bar{L}^{(i)}$, $i > 1$, be a fixed component of Σ_2. In fact we have $(C_1 \cdot L_1) = ((C_2 + \bar{L}^{(1)}) \cdot L_1) = 1$, and since $(\bar{L}^{(1)} \cdot L_1) > 0$ (the curve $L_1 + \cdots + L_s$ is connected), and $(C_2 \cdot L_1)$ cannot be negative, it follows that $(C_2 \cdot L_1) = 0$ and incidentally that

$$(\bar{L}^{(1)} \cdot L_1) = 1. \tag{7}$$

If any component of $\bar{L}^{(i)}$, $i > 1$, were a fixed component of Σ_2, then the whole curve $\bar{L}^{(i)}$ would be a fixed component of Σ_2, since it is a fundamental curve of Σ_1. We would have then $C_2 = \bar{L}^{(i)} + \bar{C}_2$ and $(\bar{C}_2 \cdot L_1) = -(\bar{L}^{(i)} \cdot L_1) < 0$, and this is impossible since L_1 is not a fixed component of Σ_2.

Thus the only possible fixed components of Σ_2 are either the irreducible

components of $\bar{L}^{(1)}$ or irreducible curves which are not components of L. We can, however, avoid this second possibility by replacing the system Σ_1 by the system $\Sigma^{(1)}$, which corresponds on F to the system cut out on $\bar{F}$ by the hyperquadrics (instead of by the hyperplanes) passing through the fundamental point O. This system $\Sigma^{(1)}$ contains all the reducible curves $C + C_1$, and it is obvious that the preceding conclusions apply as well to $\Sigma^{(1)}$ as to Σ_1. The curves of $\Sigma^{(1)}$ which contain $\bar{L}^{(1)}$ as a fixed component do not automatically possess other fixed components which are not components of L, since among these curves we find in particular the curves $L + 2C_1$. We conclude that the system $\Sigma^{(1)}$, just defined, possesses a *total* fundamental curve of the type

$$(8) \qquad L^{(1)} = k_2^{(1)} L_2 + k_3^{(1)} L_3 + \cdots + k_{s'}^{(1)} L_{s'} \qquad (s' \leq s),$$

where $L_2 + \cdots + L_{s'} = \bar{L}^{(1)}$ and where $k_2^{(1)}, \cdots, k_{s'}^{(1)}$ are positive integers. In a similar manner the remaining maximal connected components $\bar{L}^{(2)}, \cdots, \bar{L}^{(\sigma)}$ of $L_2 + \cdots + L_s$ give rise to total fundamental curves $L^{(2)}, \cdots, L^{(\sigma)}$ of $\Sigma^{(1)}$, where $L^{(i)}$ is made up of the irreducible components of $\bar{L}^{(i)}$, each counted to a proper positive multiplicity.

We have seen before that the curves C_1 of Σ_1 meet L_1 in one variable point. These curves correspond to the hyperplane sections of $\bar{F}$ on O. The curves C_1 passing through a fixed point of L_1 form a subsystem of Σ_1 of effective degree less than that of Σ_1, and hence correspond to the sections of $\bar{F}$ by hyperplanes through O and through another point O_1 of $\bar{F}$, at finite distance or infinitely near O. However, the point O_1 cannot be at a finite distance from O, since L does not carry fundamental points of the birational transformation between F and $\bar{F}$ and therefore to *every* point of the curve L there corresponds an unique point of $\bar{F}$. This point must coincide with(or be infinitely near) O, since the curve L is carried by the transformation into the point O. It follows that the curves C_1 on a fixed point of L_1 correspond to the hyperplane sections of $\bar{F}$ passing through O and touching there a fixed tangent line of $\bar{F}$. Since the curves C_1 meet L_1 in only *one* point, which varies as C_1 varies in Σ_1, it follows that there is a $(1, 1)$ correspondence between the points of L_1 and the tangential directions of $\bar{F}$ at O. The irreducible component L_1 of L thus corresponds to the first order neighborhood of the point O on $\bar{F}$. Incidentally this shows that L_1 *is a rational curve.*

We can be more precise and show that not only is L_1 a rational curve but that L_1 is free from multiple points, i. e. $p_1 = [L_1] = 0$. The system Σ_1 cuts out on L_1 a g_1^1. If L_1 had a point of multiplicity > 1, the subsystem of Σ_1 through this point would then contain L_1 as a fixed part, and this is impossible, since L_1 is not a fundamental curve of Σ_1. This property extends

similarly to the curves L_i $(i=2,\cdots,s)$. In fact, it will be shown in the next section that any curve L_{i+1} $(i=1,2,\cdots,s-1)$ plays the rôle of L_1 for a conveniently defined exceptional curve $L^{(i)}$ of the first kind [formula (16)]. Hence $p_i=0$ $(i=1,2,\cdots,s)$, and (6) becomes

$$\Sigma k_i[(L_i^2)+2]=1. \tag{6'}$$

The above considerations apply to the system $\Sigma^{(1)}$ as well as to Σ_1. Let $C^{(1)}$ denote a generic curve of $\Sigma^{(1)}$. We have $C^{(1)}\equiv C+C_1$ and hence the following relations, similar to the relations (3), hold:

$$(C^{(1)}\cdot L_1)=1,\quad (C^{(1)}\cdot L_i)=0 \qquad (i=2,\cdots,s). \tag{9}$$

The curves $C^{(1)}$ meet L_1 in one variable point, and those which pass through a fixed point of L_1 correspond to the sections of $\bar{F}$ by hyperquadrics passing through O and touching there a fixed tangent line of $\bar{F}$. Let us now consider the curve $L^{(1)}$, given by (8). In view of (7), $L^{(1)}$ meets L_1 in one point, say $P^{(1)}$. Let $O_1^{(1)}$ be the point of $\bar{F}$, infinitely near O, which corresponds to $P^{(1)}$. The curves $C^{(1)}$ of $\Sigma^{(1)}$, which are constrained to pass through $P^{(1)}$, necessarily contain the whole curve $L^{(1)}$, since $L^{(1)}$ is a fundamental curve of $\Sigma^{(1)}$. The residual system, which we shall denote by $\Sigma_1^{(1)}$, is irreducible, since it does not possess fixed components ($L^{(1)}$ is a total fundamental curve of $\Sigma^{(1)}$) and is not composed of the curves of a pencil (the system $\Sigma_1^{(1)}$ contains partially the system Σ). Moreover $\Sigma_1^{(1)}$ is of effective degree one less than that of $\Sigma^{(1)}$, since the curves of $\Sigma_1^{(1)}$ correspond to the sections of $\bar{F}$ by the hyperquadrics through O and $O_1^{(1)}$. Hence *$L^{(1)}$ is itself an exceptional curve of the first kind.* In the birational correspondence between F and $\bar{F}$ to the curve $L^{(1)}$ corresponds the point $O_1^{(1)}$ infinitely near O. The exact nature of this correspondence is yet to be investigated.

In a similar manner it can be proved that the other total fundamental curves $L^{(2)}, L^{(3)},\cdots, L^{(\sigma)}$ of $\Sigma^{(1)}$ are exceptional curves of the first kind and that they correspond to points $O_1^{(2)}, O_1^{(3)},\cdots, O_1^{(\sigma)}$ on $\bar{F}$, infinitely near and in the first neighborhood of O.

The points $O_1^{(1)}, O_1^{(2)},\cdots, O_1^{(\sigma)}$ are distinct. In fact, the point $O_1^{(\alpha)}$ corresponds, in the $(1,1)$ correspondence between the directions on $\bar{F}$ about O and the points of L_1, to the intersection $P^{(\alpha)}$ of $\bar{L}^{(\alpha)}$ or, what is the same, of $L^{(\alpha)}$ with L_1. These intersections $P^{(1)}, P^{(2)},\cdots, P^{(\sigma)}$ are distinct, since the curves $\bar{L}^{(1)}, \bar{L}^{(2)},\cdots, \bar{L}^{(\sigma)}$ have two by two no points in common.

We now prove that (7) can be replaced by the stronger relation:

$$(L^{(1)}\cdot L_1)=1. \tag{10}$$

In fact, if $C_1^{(1)}$ denotes a generic curve of $\Sigma_1^{(1)}$, so that $C^{(1)} \equiv C_1^{(1)} + L^{(1)}$, then it follows from (9) that $(L_1 \cdot C_1^{(1)}) = (L_1 \cdot C^{(1)}) - (L_1 \cdot L^{(1)}) = 1 - (L_1 \cdot L^{(1)})$. Since $(L_1 \cdot C_1^{(1)}) \geq 0$ and $(L_1 \cdot L^{(1)}) > 0$, (10) follows. We also have incidentally

$$(C_1^{(1)} \cdot L_1) = 0. \tag{10'}$$

Similar relations hold for $L^{(2)}, L^{(3)}, \cdots, L^{(\sigma)}$.

4. *The successive fundamental points $O, O_1, \cdots, O_{s-1}$ of T.* The points $O_1^{(1)}, O_1^{(2)}, \cdots, O_1^{(\sigma)}$, defined in the preceding section, are clearly to be considered as fundamental points, infinitely near O, of the birational transformation between F and $\bar{F}$. The considerations of the preceding section also show that the presence of two or more distinct fundamental points on $\bar{F}$ in the first neighborhood of O corresponds to the case in which the curve $L_2 + \cdots + L_s$ is not connected and consists of two or more maximal connected components. If then we wish to consider first the simplest case *in which there is only one fundamental point infinitely near and immediately following O,* we must assume that $L_2 + \cdots + L_s$ *is connected.* Let us make this assumption. Then the system $\Sigma^{(1)}$ possesses a total fundamental curve $L^{(1)}$ of the type:

$$L^{(1)} = k_2^{(1)} L_2 + k_3^{(1)} L_3 + \cdots + k_s^{(1)} L_s, \tag{11}$$

where $k_2^{(1)}, k_3^{(1)}, \cdots, k_s^{(1)}$ are positive integers. The fundamental point on $\bar{F}$, infinitely near O, will be denoted by O_1. We may apply to the system $\Sigma^{(1)}$ and to the exceptional curve $L^{(1)}$ the results derived in section 2 for Σ and L. Thus, choosing properly our notation, we have the following relations, analogous to the relations (3), (3') and (4):

$$(C_1^{(1)} \cdot L_2) = 1, \qquad (C_1^{(1)} \cdot L_i) = 0, \qquad (i = 3, \cdots, s); \tag{12}$$

$$k_2^{(1)} = 1; \tag{12'}$$

$$(L^{(1)} \cdot L_2) = -1; \quad (L^{(1)} \cdot L_i) = 0, \qquad (i = 3, \cdots, s). \tag{13}$$

The system $\Sigma^{(1)}$ corresponds to the sections of $\bar{F}$ by the hyperquadrics on O, and the system $\Sigma_1^{(1)}$ corresponds to the sections of $\bar{F}$ by the hyperquadrics on O and O_1. Let us transform birationally $\bar{F}$ into a surface $\bar{F}_1$ by referring the above hyperquadrics on O to the hyperplanes of a linear space. There arises a birational transformation T_1 between F and $\bar{F}_1$ in which to the system $\Sigma^{(1)}$ corresponds the system of hyperplane sections of $\bar{F}_1$. In the transformation between $\bar{F}$ and $\bar{F}_1$, which is locally quadratic at O, the first neighborhood of O is spread out into the points of a straight line; to the point

(direction) O_1 there will correspond a point O'_1 of this line, and it is clear that the system $\Sigma_1^{(1)}$ goes into the system of hyperplane sections of $\bar{F}_1$ through O'_1. For the birational transformation T_1 between F and $\bar{F}_1$, the system $\Sigma^{(1)}$ and the exceptional curve $L^{(1)}$ play the same rôle as Σ and L played in the transformation T between F and $\bar{F}$. If $s > 2$, the transformation T_1 will necessarily possess on $\bar{F}_1$ fundamental points infinitely near O'_1 and hence the transformation T will possess on $\bar{F}$ fundamental points in the second order neighborhood of O. The number of these fundamental points equals the number of maximal connected components of the curve $L_3 + \cdots + L_s$. Again, if we wish to consider the simplest case in which T possesses on $\bar{F}$ only one fundamental point O_2 in the second order neighborhood of O, we must assume that $L_3 + \cdots + L_s$ is connected.

Considering the system $\Sigma^{(2)}$ which corresponds on F to the sections of $\bar{F}_1$ by the hyperquadrics on O'_1 (this system contains all reducible curves $C^{(1)} + C_1^{(1)}$) and assuming that $L_3 + \cdots + L_s$ is connected, we have that $\Sigma^{(2)}$ possesses a total fundamental curve of the type

$$L^{(2)} = k_3^{(2)} L_3 + \cdots + k_s^{(2)} L_s,$$

and that $L^{(2)}$ is an exceptional curve of the first kind. If $\Sigma_1^{(2)}$ denotes the residual system of $\Sigma^{(2)}$ with respect to $L^{(2)}$ and if $C^{(2)}$ and $C_1^{(2)}$ denote generic curves of $\Sigma^{(2)}$ and $\Sigma_1^{(2)}$ respectively, then, with a proper choice of notation, we have the following relations similar to (12), (12′) and (13):

$$(14) \qquad (C_1^{(2)} \cdot L_3) = 1, \qquad (C_1^{(2)} \cdot L_i) = 0 \qquad (i = 4, \cdots, s);$$

$$(14') \qquad k_3^{(2)} = 1;$$

$$(15) \qquad (L^{(2)} \cdot L_3) = -1, \quad (L^{(2)} \cdot L_i) = 0 \qquad (i = 4, \cdots, s).$$

If $s > 3$ the transformation T between F and $\bar{F}$ possesses on $\bar{F}$ fundamental points in the third order neighborhood of O (arising from fundamental points of T_1 on $\bar{F}_1$ in the second order neighborhood of O'_1) and there is but one such fundamental point if and only if $L_4 + \cdots + L_s$ is connected.

The general procedure is now straight-forward and serves to define the fundamental points of T on $\bar{F}$ in the 1st, 2nd, $\cdots$, $(s-1)$-th order neighborhoods of O. *We assume at present that each fundamental point O_i in the i-th neighborhood of O $(i < s-1)$ is followed by just one fundamental point O_{i+1} in the neighborhood of order $i+1$.* With these assumptions, T possesses on $\bar{F}$ in addition to O other $s-1$ successive fundamental points $O_1, \cdots, O_{s-1}$, infinitely near O_1, *lying on one irreducible algebroid branch of origin O.*

We may define by induction the systems $\Sigma^{(i)}$, $\Sigma_1^{(i)}$, the exceptional curves

$L^{(i)}$ $(i=0,1,\cdots,s-1)$ and state their properties as follows. $L^{(i)}$ is a total fundamental curve of $\Sigma^{(i)}$ and we have

$$L^{(i)} = k_{i+1}^{(i)} L_{i+1} + k_{i+2}^{(i)} L_{i+2} + \cdots + k_s^{(i)} L_s, \tag{16}$$

where the k's are positive integers. $\Sigma_1^{(i)}$ is the residual system of $\Sigma^{(i)}$ (of dimension one less than $\Sigma^{(i)}$) with respect to $L^{(i)}$. We have the following relations, similar to the relations (14), (14′), (15):

$$(C_1^{(i)} \cdot L_{i+1}) = 1, \quad (C_1^{(i)} \cdot L_j) = 0 \qquad (j = i+2, \cdots, s); \tag{17}$$

$$k_{i+1}^{(i)} = 1; \tag{17′}$$

$$(L^{(i)} \cdot L_{i+1}) = -1; \quad (L^{(i)} \cdot L_j) = 0 \qquad (j = i+2, \cdots, s). \tag{18}$$

Here $C^{(i)}$ and $C_1^{(i)}$ denote total curves of $\Sigma^{(i)}$ and $\Sigma_1^{(i)}$ respectively. *The curve $L_{i+2} + \cdots + L_s$ is connected.* The system $\Sigma^{(i+1)}$ is the minimum linear system containing as total curves all reducible curves $C^{(i)} + C_1^{(i)}$.

Let $\bar{F}_i$ be the birational transform of the surface F obtained by referring the curves of $\Sigma^{(i)}$ to the hyperplanes of a linear space, and let T_i denote the birational transformation between F and $\bar{F}_i$ $(\bar{F}_0 = F,\ \Sigma^{(0)} = \Sigma,\ T_0 = T)$. The system $\Sigma_1^{(i)}$ corresponds by T_i to the system of hyperplane sections of $\bar{F}_i$ on a fixed point $O_i^{(i)}$, while the curves of $\Sigma^{(i+1)}$ correspond to the sections of $\bar{F}_i$ by the hyperquadrics through $O_i^{(i)}$. It is clear that $\bar{F}_{i+1}$ is a birational transform of $\bar{F}_i$, obtained by referring the sections of $\bar{F}_i$ by the hyperquadrics through $O_i^{(i)}$ to the hyperplane sections of $\bar{F}_{i+1}$. In this transformation, which is locally quadratic at $O_i^{(i)}$, the point $O_{i+1}^{(i+1)}$, corresponds to a direction about $O_i^{(i)}$. The points $O_0^{(0)} = O, O_1^{(1)}, O_2^{(2)}, \cdots, O_{s-1}^{(s-1)}$ are the transforms of the s fundamental points $O, O_1, \cdots, O_{s-1}$ on $\bar{F}$ by the above successive locally quadratic transformations and serve to define the position of the fundamental points $O, O_1, \cdots, O_{s-1}$ on $\bar{F}$ on algebroid branches: every branch on $\bar{F}$ passing through $O, O_1, \cdots, O_i$ corresponds to a branch on $\bar{F}_i$ passing through $O_i^{(i)}$.

The following intersection formulas will be useful in the sequel and are obtained by applying the formulas (10) and (10′) to the exceptional curves $L^{(i)}$:

$$(L^{(i)} \cdot L_i) = 1; \tag{19}$$

$$(C_1^{(i)} \cdot L_i) = 0 \qquad (i = 1, \cdots, s-1). \tag{19′}$$

Remark. Our notation implies a definite ordering of the irreducible components $L_1, L_2, \cdots, L_s$ of L. From (19) it follows that *each curve L_i*

intersects *one and only one of the curves* $L_{i+1}, \cdots, L_s$. If $L_{i+\alpha}$ $(\alpha > 0)$ is the curve which L_i intersects, then $(L_i \cdot L_{i+\alpha}) = 1$ and $L_{i+\alpha}$ must be a simple component of $L^{(i)}$. *It does not, however, necessarily coincide with the simple component* L_{i+1}, *as we shall see in the sequel.* At any rate, all the intersection numbers $(L_i \cdot L_j)$, $i \neq j$, are either 0 or 1.

5. *The correspondence between* L_i *and the immediate neighborhood of* O_{i-1}. It has been proved in section 3 that, in the birational correspondence T between F and $\bar{F}$, the irreducible component L_1 of L corresponds to the first order neighborhood of the fundamental point O on $\bar{F}$. In a similar manner, in the birational correspondence T_{i-1} between F and $\bar{F}_{i-1}$ the irreducible component L_i of $L^{(i-1)}$ corresponds to the first order neighborhood of the fundamental point $O_{i-1}^{(i-1)}$ on $\bar{F}_{i-1}$. Hence, to a branch δ_{i-1} on $\bar{F}_{i-1}$ of origin $O_{i-1}^{(i-1)}$ and *possessing at* $O_{i-1}^{(i-1)}$ *a principal tangent line distinct from the singular direction at* $O_{i-1}^{(i-1)}$ (that direction to which there corresponds the point $O_i^{(i)}$ on $\bar{F}_i$), there corresponds a branch γ_{i-1} on F meeting L_i at point P_i. The branch δ_{i-1} arises, by the sequence of locally quadratic transformations applied to $\bar{F}, \bar{F}_1, \cdots$, from a branch $\bar{\gamma}_{i-1}$ passing through $O, O_1, \cdots, O_{i-1}$ *but not through* O_i. Hence we have the following

THEOREM 2. *To a branch* $\bar{\gamma}_{i-1}$ *on* $\bar{F}$, *passing through* $O, O_1, \cdots, O_{i-1}$ *but not through* O_i, *there corresponds a branch* γ_{i-1} *on* F *meeting* L_i. *As we approach on* $\bar{F}$ *the point* O *along such a branch* $\bar{\gamma}_{i-1}$ *the homologous point on* F *approaches a point* P_i *on* L_i. *This point* P_i *varies as the point on* $\bar{\gamma}_{i-1}$ *which immediately follows* O_{i-1} *varies.*

It will be noted that an apparent discontinuity arises as $\bar{\gamma}_{i-1}$ approaches a branch $\bar{\gamma}_i$ passing through $O, O_1, \cdots, O_{i-1}$ *and* O_i. The branch γ_i, which corresponds to $\bar{\gamma}_i$ on $\bar{F}$, is a branch meeting L_{i+1} but not necessarily L_i. The real situation is the following: as $\bar{\gamma}_{i-1}$ approaches a branch $\bar{\gamma}_i$, the corresponding branch γ_{i-1} on F approaches a locus which degenerates into the curve $L^{(i)}$ and into the branch γ_i.

Another paradoxical circumstance is the following. We have seen that there is a $(1, 1)$ correspondence between the points of L_i and the points on $\bar{F}$ infinitely near and immediately following O_{i-1}. Among these points there is the fundamental point O_i, and it seems natural to expect that the corresponding point on L_i is a common point of L_i and L_{i+1}. However, this is not always the case, *as it may well happen that* L_i *and* L_{i+1} *do not intersect at all.* We shall deal with the general case in the next section. Here an example may suffice. Let $s = 3$ and let the fundamental points O, O_1, O_2 on $\bar{F}$ lie on a

cuspidal branch. The general results of the following section show that in this case we have: $L = L_1 + L_2 + 2L_3$, $(L_1 \cdot L_2) = 0$, $(L_1 \cdot L_3) = 1$, $(L_2 \cdot L_3) = 1$ and incidentally $(L_1^2) = -3$, $(L_2^2) = -2$, $(L_3^3) = -1$. Thus the curve L_1 does not meet L_2, although there is a point on L_1 which corresponds formally to the direction $O\,O_1$. This point is the intersection of L_1 and L_3. If we consider a branch on F passing through this point and if we assume for simplicity that this branch is linear, we find that the corresponding branch on $\bar{F}$ passes *triply* through O and *simply* through O_1 and O_2.

6. *The geometric significance of the coefficients* $k_1^{(j)}$. A branch $\bar{\gamma}_{i-1}$ on $\bar{F}$ of *smallest order*, passing through $O, O_1, \cdots, O_{i-1}$, arises from a linear branch δ_{i-1} on $\bar{F}_{i-1}$, of origin $O_{i-1}^{(i-1)}$ and possessing at $O_{i-1}^{(i-1)}$ a generic tangent line. To such a branch there corresponds on F a *linear* branch γ_{i-1} of origin P_{i-1}, a generic point of L_i and therefore not on L_j, $j \neq i$.* It is well known from the theory of singularities that branches $\bar{\gamma}_{i-1}$ of lowest order are also characterized by the fact that such a branch $\bar{\gamma}_{i-1}$ passes *simply* through O_{i-1} and that the points which follow O_{i-1} on $\bar{\gamma}_{i-1}$ are all *free points*. ([3], pp. 365-366, 372).

The orders of multiplicity with which $\bar{\gamma}_{i-1}$ passes through the points $O, O_1, \cdots, O_{i-1}$ can be obtained as follows. The order of the multiple point O_{j-1} $(j \leq i)$ equals the multiplicity with which the branch δ, which corresponds on the surface $\bar{F}_{j-1}$ to $\bar{\gamma}_{i-1}$, passes through the point $O_{j-1}^{(j-1)}$, its origin. We may assume that δ is the branch of some irreducible algebraic curve $\bar{D}$ on $\bar{F}_{j-1}$ and that $\bar{D}$ does not possess other branches of origin $O_{j-1}^{(j-1)}$, distinct from δ. Then we may calculate the multiplicity of $\bar{D}$ (i. e. of δ) at $O_{j-1}^{(j-1)}$ as the difference between the number of intersections of a variable hyperplane section of $\bar{F}_{j-1}$ with $\bar{D}$ and the number of variable intersections with $\bar{D}$ of a hyperplane section through $O_{j-1}^{(j-1)}$. Going back to the surface F we see that the required multiplicity equals $(C^{(j-1)} \cdot D) - (C_1^{(j-1)} \cdot D)$, where D on F is the transform of $\bar{D}$ on $\bar{F}_{j-1}$. Now $C^{(j-1)} = C_1^{(j-1)} + L^{(j-1)}$ and hence $(C^{(j-1)} \cdot D) - (C_1^{(j-1)} \cdot D) = (L^{(j-1)} \cdot D) = k_i^{(j-1)}$, since the branch γ_{i-1}, and hence also D, has a simple intersection with the curve L_i, does not meet any other irreducible component of $L^{(j-1)}$, and since L_i is a $k_i^{(j-1)}$-fold component of $L^{(j-1)}$. We thus have the following

THEOREM 3. *The fundamental points* $O, O_1, \cdots, O_{i-1}$ $(i \leq s)$ *on* $\bar{F}$ *lie*

* The truth of this statement follows immediately if we recall that to the sections of $\bar{F}_{i-1}$ by hyperplanes through $O_{i-1}^{(i-1)}$ there correspond on F the curves $C_1^{(i-1)}$ of the system $\Sigma_1^{(i-1)}$ and that by (17) $(C_1^{(i-1)} \cdot L_i) = 1$.

on a branch $\bar{\gamma}_{i-1}$ of lowest order, passing through these points with the multiplicities $k_i, k_i^{(1)}, k_i^{(2)}, \cdots, k_i^{(t-1)} = 1$.

As a corollary we have the following inequalities:

$$k_i \geq k_i^{(1)} \geq k_i^{(2)} \geq \cdots \geq 1. \tag{20}$$

In particular the branch of lowest order passing through all the s fundamental points $O, O_1, \cdots, O_{s-1}$ has at these points the multiplicities $k_s, k_s^{(1)}, k_s^{(2)}, \cdots, k_s^{(s-1)} = 1$ respectively.

The branch $\bar{\gamma}_i$ passes through $O, O_1, \cdots, O_{i-1}$ and in addition through O_i. The multiplicity of any point O_j, $j \leq i-1$, on this branch cannot be less than its multiplicity for the branch $\bar{\gamma}_{i-1}$ of lowest order passing through $O, O_1, \cdots, O_{i-1}$. Hence we also have the following inequalities:

$$k_j^{(j-1)} \leq k_{j+1}^{(j-1)} \leq k_{j+2}^{(j-1)} \leq \cdots \leq k_s^{(j-1)}. \tag{21}$$

7. *The determination of the intersection numbers* $(L_i \cdot L_j)$. In his theory of plane singularities, Enriques ([3], p. 381) has introduced the notion of *proximate points* on an algebroid branch. Let $O, O_1, \cdots, O_{s-1}$ be a sequence of infinitely near points on an algebroid branch γ, and let ν_i be the multiplicity of O_i on γ. We have for any i, $\nu_i \geqq \nu_{i+1}$. If $\nu_i = \nu_{i+1}$, then the set of proximate points of O_i on γ consists by definition of the single point O_{i+1}. If $\nu_i > \nu_{i+1}$ and if $\nu_i = h\nu_{i+1} + \nu'$ $(0 \leq \nu' < \nu_{i+1})$, then it can be proved ([3], p. 381) that O_i is followed immediately on γ by h successive ν_{i+1}-fold points $O_{i+1}, O_{i+2}, \cdots, O_{i+h}$ and by one ν'-fold point O_{i+h+1} $(i+h+1 \leq s-1)$, if $\nu' > 0$. These $h+1$ points constitute the set of proximate points of O_i. If $\nu' = 0$, the set of proximate points of O_i consists of h points $O_{i+1}, \cdots, O_{i+h}$. From the above definition one deduces the following fundamental property of the set of proximate points of a given point: *the multiplicity of O_i on γ equals the sum of the multiplicities of its proximate points.*

We shall have occasion to use the following property of proximate points:

THEOREM 4. *If O_{i+a} is in the set of proximate points of O_i on a given branch γ, then it also belongs to the set of proximate points of O_i on any other branch passing through O_{i+a} (and hence also through O_i).*

This theorem is implicitly contained in Enriques and follows from the construction of the branches of lowest order passing through a given set of successive infinitely near points $O, O_1, \cdots, O_i, O_{i+1}, \cdots, O_{s-1}$ on a given branch γ. If $O_{i+1}, \cdots, O_{i+\beta}$ is the set of proximate points of O_i on γ, it is found that on a branch of lowest order passing through $O, O_1, \cdots, O_{s-1}$ the

set of proximate points of O_i consists of the set $O_{i+1}, \cdots, O_{i+\beta}$ if $s-1 > i+\beta$, and includes the set $O_{i+1}, \cdots, O_{s-1}$ if $s-1 \leq i+\beta$ ([4], Chap. I, 2). From this the above theorem follows immediately. We shall now prove the following

THEOREM 5. *Let γ_{s-1} be the branch of lowest order which passes through the infinitely near fundamental points $O, O_1, \cdots, O_{s-1}$ on $\bar{F}$ and let the set of proximate points of O_{i-1} on γ_{s-1} $(i < s)$ consist of α_i points $O_i, O_{i+1}, \cdots, O_{i+\alpha_i-1}$.* Then $(L_i \cdot L_{i+\alpha_i}) = 1$, $(L_i \cdot L_j) = 0$ if $j > i$ and $j \neq i+\alpha_i$. Moreover the virtual degree (L_i^2) of L_i equals $-(1+\alpha_i)$.*

For the proof we observe first of all that it is sufficient to prove the theorem for $i = 1$ $(O_{i-1} = O,\ L_i = L_1)$, for then we may apply the result to any exceptional curve $L^{(i-1)}$ and to the sequence of fundamental points $O_{i-1}^{(i-1)}, O_i^{(i-1)}, \cdots, O_{s-1}^{(i-1)}$ on $\bar{F}_{i-1}$ of the birational correspondence between F and $\bar{F}_{i-1}$. These points lie on a branch of lowest order, which is the transform of the branch γ_{s-1}, and the proximate points of $O_{i-1}^{(i-1)}$ are clearly the points $O_i^{(i-1)}, \cdots, O_{i+\alpha_i-1}^{(i-1)}$.

We first prove the theorem when $\alpha = 1$, i. e. when the set of proximate points of O consists of the single point O_1. By Theorem 3 we have then $k_s = k_s^{(1)}$. Applying Theorem 4, we see that the set of proximate points of O on any branch γ_{i-1} of lowest order passing through $O, O_1, \cdots, O_{i-1}$ will consist of the single point O_1, and hence, again by Theorem 3, we have $k_i^{(1)} = k_i$ $(i = 2, 3, \cdots, s)$, i. e. $L = L_1 + L^{(1)}$. By (18) we have $(L \cdot L_2) = 0$, $(L^{(1)} \cdot L_2) = -1$, hence $(L_1 \cdot L_2) = 1$ and this proves the first part of the theorem. We also have $(L \cdot L_1) = (L_1^2) + (L^{(1)} \cdot L_1)$ and by (18) and (19) $(L \cdot L_1) = -1$, $(L^{(1)} \cdot L_1) = 1$. Consequently $(L_1^2) = -2$, q. e. d.

By section 4, Remark, L_1 meets only one of the curves $L_2, \cdots, L_s$ and the corresponding intersection number is 1. To prove our theorem for α arbitrary, $\alpha > 1$, we shall assume that $(L_1 \cdot L_{\alpha+1}) = 1$ and we shall show that then the set of proximate points of O on the branch γ_{s-1} consists of the points $O_1, \cdots, O_\alpha$ and that $(L_1^2) = -(1+\alpha)$.

We have $(L^{(1)} \cdot L_1) = k_{\alpha+1}^{(1)}$ and hence by (19) $k_{\alpha+1}^{(1)} = 1$. Using the inequalities (20) we find $k_{\alpha+1}^{(2)} = k_{\alpha+1}^{(3)} = \cdots = k_{\alpha+1}^{(\alpha)} = 1$, and from this it follows, in view of the inequalities (21), that all the coefficients $k_i^{(j)}$ $(j = 1, 2, \cdots, \alpha,\ i = j+1, \cdots, \alpha+1)$, are 1. For the sake of clearness we write the simplified expression of the curves $L, L^{(1)}, \cdots, L^{(\alpha)}$:

---

* Here necessarily $i + \alpha_i - 1 \leq s - 1$ because on a branch of lowest order through a given sequence of infinitely near points $O, O_1, \cdots, O_{s-1}$, the set of proximate points of any point of the sequence except the last point O_{s-1} is contained in the given sequence.

$$
\begin{aligned}
L &= L_1 + L_2 + k_3 L_3 + \cdots + k_a L_a + k_{a+1} L_{a+1} + k_{a+2}\ L_{a+2} + \cdots \\
L^{(1)} &= \quad L_2 + \ L_3 + \cdots + \ L_a + \quad L_{a+1} + k^{(1)}_{a+2}\ L_{a+2} + \cdots \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
L^{(a-1)} &= \qquad\qquad\qquad\qquad\quad L_a + \quad L_{a+1} + k^{(a-1)}_{a+2} L_{a+2} + \cdots \\
L^{(a)} &= \qquad\qquad\qquad\qquad\qquad\quad L_{a+1} + k^{(a)}_{a+2}\ L_{a+2} + \cdots .
\end{aligned}
$$

We have replaced k_2 by 1, since k_2 and $k_2^{(1)}$ $(=1)$ give the multiplicities of O and O_1 on the branch of lowest order passing through O and O_1, and hence clearly $k_2 = 1$. From Theorem 3 it follows that the branch γ_a of lowest order passing through $O, O_1, \cdots, O_a$ has at these points the multiplicities $k_{a+1}, 1, \cdots, 1$. Consequently on this branch the set of proximate points of O_i, $1 \leq i \leq \alpha - 1$, consists of the single point O_{i+1}. It follows by Theorem 4 that also on γ_{s-1} the set of proximate points of O_i, where now $1 \leq i \leq \alpha - 2$, consists of the single point O_{i+1}. Since we have already proved our theorem for the case $\alpha = 1$, we conclude that the following relations hold:

$$(22) \qquad (L_i \cdot L_{i+1}) = 1$$

$$(22') \qquad (L_i^2) = -2 \qquad\qquad (i = 2, 3, \cdots, \alpha - 1).$$

By (18) we have $(L \cdot L_1) = -1$ and $(L \cdot L_i) = 0$, $i > 1$. Using the relations (22), (22′) and recalling that by section 4, Remark, each curve L_i intersects only one of the components L_j, $j > i$, we find

$$(23) \qquad (L_1^2) + k_{a+1} = -1;$$

$$(23') \qquad -2k_2 + k_3 = 0;$$

$$(23'') \qquad k_{i-1} - 2k_i + k_{i+1} = 0 \qquad\qquad (i = 3, \cdots, \alpha - 1).$$

Since $k_1 = k_2 = 1$ the relations (23′) and (23″) give the following values for $k_3, k_4, \cdots, k_a$:

$$(24) \qquad k_3 = 2, k_4 = 3, \cdots, k_a = \alpha - 1.$$

The curve L_a intersects one (and only one) of the curves $L_{a+1}, L_{a+2}, \cdots$. Let $(L_a \cdot L_{a+j}) = 1$, $j \geq 1$. From $(L^{(i)} \cdot L_a) = 0$ $(i = 1, 2, \cdots, \alpha - 2)$, $(L^{(a-1)} \cdot L_a) = -1$ and $(L^{(a)} \cdot L_a) = 1$ we derive the following relations:

$$(25) \qquad 1 + (L_a^2) + k^{(i)}_{a+j} = 0 \qquad\qquad (i = 1, 2, \cdots, \alpha - 1);$$

$$(25') \qquad k^{(a)}_{a+j} = 1.$$

From these relations it follows that $k^{(1)}_{a+j} = k^{(2)}_{a+j} = \cdots = k^{(a-1)}_{a+j}$ and $k^{(a)}_{a+j} = k^{(a+1)}_{a+j} = \cdots = k^{(a+j-1)}_{a+j} = 1$ by (20). Hence the branch γ_{a+j-1} of lowest order passing through the points $O, O_1, \cdots, O_{a+j-1}$ has at these points the following multiplicities: $O^{ka+j}, O_1{}^k, O_2{}^k, \cdots, O^k{}_{a-1}, O_a{}^1, \cdots, O^1{}_{a+j-1}$, where we have put $k = k^{(1)}_{a+j}$. We also have $(L \cdot L_a) = 0$ and hence from (24)

$$(\alpha-2)+(\alpha-1)(L_\alpha^2)+k_{\alpha+j}=0.$$

Substituting for (L_α^2) the value $-k-1$ from (25), we find

$$k_{\alpha+j}=(\alpha-1)k+1.$$

This value of $k_{\alpha+j}$ shows that on the above branch $\gamma_{\alpha+j-1}$ the set of proximate points of O consists of the points $O_1, O_2, \cdots, O_\alpha$. By Theorem 4 it follows then that also on the branch γ_{s-1} the points $O_1, \cdots, O_\alpha$ constitute the set of proximate points of O, and this proves the first part of the theorem.

We have already observed that the branch γ_α of lowest order passing through the points $O, O_1, \cdots, O_\alpha$ possesses at these points the following multiplicities: $O^{k_{\alpha+1}}, O_1^1, \cdots, O_\alpha^1$. Also on this branch the points $O_1, \cdots, O_\alpha$ must constitute the set of proximate points of O. Hence $k_{\alpha+1}=\alpha$ and by (23), $(L_1^2)=-(1+\alpha)$, q. e. d.

Remark 1. By Theorem 1, the curve L_s is of virtual degree -1 because $L_s=L^{(s)}$, i. e. L_s is itself an exceptional curve of the first kind.

Remark 2. The preceding theorem shows that the intersection numbers $(L_i \cdot L_j)$ are completely determined by the characters of the branch γ_{s-1} of lowest order passing through the infinitely near fundamental points $O, \cdots, O_{s-1}$. Conversely, it is not difficult to show that *the intersection numbers* $(L_i \cdot L_j)$ *completely determine the characters of the branch* γ_{s-1}, i. e. *the multiplicities of the points* $O, O_1, \cdots, O_{s-1}$ *on this branch.* In fact, $-(L_i^2)-1$ gives the number of proximate points of O_i $(i<s-1)$, and hence the multiplicity of O_i on γ_{s-1} can be found if the multiplicities of the points $O_{i+1}, O_{i+2}, \cdots, O_{s-1}$ are known, because the multiplicity of O_i equals the sum of the multiplicities of the proximate points. Since the multiplicity of O_{s-1} is 1, the multiplicities of $O_{s-2}, O_{s-3}, \cdots$ can be determined step by step.

In a similar manner the multiplicities of the points $O, O_1, \cdots, O_{i-1}$ on the branch γ_{i-1} of lowest order containing them can be determined. Hence, by Theorem 3, it follows that *the intersection numbers* $(L_i \cdot L_j)$ *determine uniquely all the integers* $k_i^{(j)}$.

Remark 3. *The determinant* $\Delta_i=|(L_\alpha \cdot L_\beta)|$ $(\alpha, \beta=i, i+1, \cdots, s)$, *equals* $(-1)^{s-i+1}$. In fact, if we multiply in Δ_1 the second column by k_2, the third column by $k_3, \cdots$, the s-th column by k_s and add to the first column, we find, in view of equations (5), $\Delta_1=-\Delta_2$. In a similar manner we find $\Delta_2=-\Delta_3=\Delta_4=\cdots=\pm\Delta_s$. Since $\Delta_s=(L_s^2)=-1$, the statement follows. The equations (5) determine the integers k_i in terms of the intersection numbers $(L_i \cdot L_j)$. This is in agreement with the preceding remark.

8. *Free points and satellites.* Our analysis of reducible exceptional curves runs parallel to the analysis of plane singularities due to Enriques. It illustrates very concretely, by means of the intersection properties of the components of the exceptional curve, the notions and conventions used in this theory, as for instance the notion of proximate points. Another concept, also due to Enriques ([3], pp. 365-366, 372), is that of *free points and satellites* in a sequence of infinitely near points on an algebroid branch. We shall now show that the knowledge of the intersection numbers $(L_i \cdot L_j)$ allows us to decide whether the fundamental point O_i on γ_{s-1}, which corresponds to a given component L_{i+1} of L, is a free point or a satellite. For this we quote a few properties relative to the classification of the points $O, O_1, \cdots, O_{s-1}$ into free points and satellites ([4], Chap. I, 2).

Let $O_{k+1}, O_{k+2}, \cdots, O_{k+l}$ be a sequence of free points such that O_k (if $O_{k+1} \neq O$) and O_{k+l+1} (if $k+l<s-1$) are satellites, and let ν_i denote the multiplicity of O_i on γ_{s-1}. Such a sequence enjoys the following characteristic properties:

(α) $\nu_{k+1} = \nu_{k+2} = \cdots = \nu_{k+l-1}$; $\nu_{k+l} < \nu_{k+l-1}$ except when $k+l=s-1$, in which case $\nu_{k+l-1} = \nu_{k+l} = 1$.

(β) If O_{k+1} is distinct from O, then $k \geq 2$ and $\nu_{k-1} = \nu_k$. Moreover if $\nu_k = \nu_{k-1} = \cdots = \nu_{k-i+1}$ and $\nu_k \neq \nu_{k-i}$, then $\nu_{k-i} = i\nu_k$.

Recalling the properties of proximate points quoted in section 7, we conclude from (α) and (β) that

(1) Every satellite is the last proximate point of at least one point preceding it.

(2) If O_k is a satellite followed by a free point O_{k+1}, then O_k is the last proximate point of two points O_{k-1} and O_{k-i} preceding it, and conversely.

(3) The last proximate point of any point O_i is never a free point followed by a satellite.

From (1), (2) and (3) there follows

THEOREM 6. *If L_{i+1} does not meet any curve L_j, $j<i+1$, then O_i is a free point followed by a satellite O_{i+1}, and conversely. If L_{i+1} meets two of the curves L_j, $j<i+1$, then O_i is a satellite followed by a free point O_{i+1}, and conversely.*

This theorem characterizes the *last* point of a sequence of satellites followed by free points and the *last* point of a sequence of free points followed by satellites and hence enables us to find out the division of the set $O, O_1, \cdots, O_{s-1}$ into free points and satellites from the intersection properties of the curves $L_1, L_2, \cdots, L_s$.

We add a few remarks concerning the $(1, 1)$ correspondence between the points of L_{i+1} and the points of the immediate neighborhood of O_i (points in the neighborhood of order $i+1$ of O). It is not difficult to show ([4], Chap. I, 2) that the simple infinity of points immediately following O_i contains *one* satellite, if O_i is a free point, and *two* satellites if O_i is itself a satellite. What point or points of L_{i+1} correspond to this satellite or these satellites? To answer this question it is necessary to observe that if $\bar{O}_{i+1}$ is a point in the immediate neighborhood of O_i, then the branch of lowest order passing through $O, O_1, \cdots, O_i, \bar{O}_{i+1}$ is of the same order as the branch γ_i of lowest order passing through $O, O_1, \cdots, O_i$ if $\bar{O}_{i+1}$ is a free point, and is of higher order than γ_i if $\bar{O}_{i+1}$ is a satellite. A branch γ_i corresponds to a *linear* branch on F, meeting L_{i+1} in a generic point. We obtain on $\bar{F}$ a branch of higher order than γ_i if and only if that linear branch intersects L_{i+1} in a point where L_{i+1} intersects some further component of L. Recalling that the fundamental point O_{i+1} corresponds formally to the intersection of L_{i+1} with a curve L_j, $j > i+1$ (section 5), we conclude as follows: if O_i is a free point and O_{i+1} is a satellite, then this satellite corresponds to the intersection of L_{i+1} with an L_j, $j > i+1$ (it should be noticed that in this case, by Theorem 6, L_{i+1} does not meet any curve L_j, $j < i+1$). If O_i and O_{i+1} are both free points, then, by Theorem 6, L_{i+1} necessarily intersects one and only one curve L_j, $j < i+1$; to this point of intersection corresponds *the* satellite which immediately follows O_i. If O_i is a satellite and O_{i+1} is a free point, then, by Theorem 6, L_{i+1} meets two curves L_j, $j < i+1$, and the two points of intersection correspond to the two satellites in the immediate neighborhood of O_i. Finally if O_i and O_{i+1} are both satellites, then by Theorem 6 L_{i+1} meets only one curve L_j, $j < i+1$, and the point of intersection corresponds to the second satellite, distinct from O_{i+1}, which immediately follows O_i.

9. *Characterization of reducible exceptional curves of the first kind.* It is known ([1]) that the conditions which characterize an *irreducible* exceptional curve L_1 of the first kind, on a surface free from singularities, are the following: $(L_1^2) = -1$, $[L_1] = 0$, these characters being evaluated on the assumption that L_1 is virtually free from base points. The curve L_1 is rational and *free from singularities*, because, if L_1 had singularities, the virtual genus $[L_1]$ would be positive.

We show that the properties of the intersection numbers $(L_i \cdot L_j)$, derived in the preceding section, together with the relations $[L_i] = 0$, proved in section 3, characterize these reducible exceptional curves of the first kind, which arise from birational transformations possessing a sequence of infinitely

near fundamental points lying on an irreducible algebroid branch. In exact terms we prove the following

THEOREM 7. *Let $L_1, L_2, \cdots, L_s$ be a set of irreducible curves, virtually free from base points and of virtual genus zero, on a surface F free from singularities, and let there exist an arrangement of these curves, for instance the one written above, such that the following conditions are satisfied:* (i) *each curve L_i meets one and only one of the curves $L_{i+1}, \cdots, L_s$ which follow L_i;* (ii) *if L_i meets $L_{i+\alpha}$ $(i < s)$, then $(L_i \cdot L_{i+\alpha}) = 1$, $(L_i^2) = -(1+\alpha)$ and $(L_{i+j} \cdot L_{i+j+1}) = 1$ $(j = 1, 2, \cdots, \alpha - 2)$;* (iii) $(L_s^2) = -1$. *Under these conditions a convenient combination $k_1L_1 + k_2L_2 + \cdots + k_sL_s$ $(k_1 = k_2 = 1)$ with positive integral coefficients is an exceptional curve of the first kind. The assumed arrangement $L_1, \cdots, L_s$ is uniquely determined and so also are the integers k_i.*

Proof. Since the theorem is true for $s = 1$ ([1]), we shall prove it by induction, assuming it to be true for $s - 1$. Since, by hypothesis, $[L_s] = 0$ and $(L_s^2) = -1$, L_s is an exceptional curve of the first kind. There exists therefore a birational transformation of F into a surface $\bar{F}$, in which to L_s there corresponds a simple point O of $\bar{F}$ and which does not possess fundamental points on L_s. It can also be assumed that the transformation possesses no fundamental points on F, that L_s is the only fundamental curve of the transformation on F and that $\bar{F}$ is free from singularities.* Let $\bar{L}_1, \bar{L}_2, \cdots, \bar{L}_{s-1}$ be the transforms of $L_1, \cdots, L_{s-1}$ on F. The virtual degree of a curve is invariant under birational transformations. However, if a given L_i meets L_s, then L_i passes through the point O *simply*, since $(L_i \cdot L_s) = 1$, and this point must be considered as an assigned base point of $\bar{L}_i$ ([4], Chap. II, 7). If we understand by $(\bar{L}_i^2)$ the virtual degree of $\bar{L}_i$ *virtually free from base points*, then $(\bar{L}_i^2) = (L_i^2)$ if $(L_i \cdot L_s) = 0$ and $(\bar{L}_i^2) = (L_i^2) + 1$ if $(L_i \cdot L_s) = 1$.

With these preliminaries we proceed to prove that the ordered sequence of curves $\bar{L}_1, \bar{L}_2, \cdots, \bar{L}_{s-1}$ on $\bar{F}$ satisfies the conditions for an exceptional curve stated in the above theorem. From (i) we deduce that $(L_{s-1} \cdot L_s) = 1$ and by (ii) $(L^2_{s-1}) = -2$. Hence $(\bar{L}^2_{s-1}) = (L^2_{s-1}) + 1 = -1$ and this shows that the condition (iii) holds. To show that the conditions (i) and (ii) are satisfied, we make the following self-evident remarks:

---

* The defining linear system Σ on F associated with the transformation, i. e. the system which goes into the system of hyperplane sections of $\bar{F}$, may be assumed to be free from base points and non-singular ([1], [4], Chap. IV, 4). If the transformation possesses on F fundamental curves other than L_s and hence fundamental points on $\bar{F}$ distinct and necessarily at a finite distance from O, these fundamental points can be transformed back into curves by means of a sufficiently general linear system on F, free from fundamental curves and free from base points outside these fundamental points.

(α) If $(L_i \cdot L_s)$ and $(L_j \cdot L_s)$ are not both 1 $(i, j < s)$, then $(\bar{L}_i \cdot \bar{L}_j) = (L_i \cdot L_j)$.

(β) If $(L_i \cdot L_s) = (L_j \cdot L_s) = 1$, then $(\bar{L}_i \cdot \bar{L}_j) = (L_i \cdot L_j) + 1$.

We consider any curve L_i, $i < s-1$. Let $(L_i \cdot L_{i+\alpha}) = 1$, $(L_i \cdot L_j) = 0$, $j > i$ and $j \neq i+\alpha$. If $i+\alpha \neq s$, then by (α) $(\bar{L}_i \cdot \bar{L}_{i+\alpha}) = 1$ and $(\bar{L}_i \cdot \bar{L}_j) = 0$ if $j > i$ and $\neq i+\alpha$. Moreover by (α) applied to the case $i = j$, it follows that $(\bar{L}_i^2) = (L_i^2) = -1-\alpha$. If $i+\alpha = s$, then by (β) (where now $j = s-1$) we have $(\bar{L}_i \cdot \bar{L}_{s-1}) = 1$, since $(L_i \cdot L_{s-1}) = 0$. Moreover by (ii) we have $(L_j \cdot L_s) = 0$ $(j = i+1, \cdots, s-2)$, and hence by (α) $(\bar{L}_i \cdot \bar{L}_j) = 0$ $(j = i+1, \cdots, s-2)$, since $(L_i \cdot L_j) = 0$. We also have by (β), when we put $i = j$, $(\bar{L}_i^2) = (L_i^2) + 1 = -(s-i) - 1 + 1 = -(s-i-1) - 1$. We have thus proved that every curve $\bar{L}_i$ $(i = 1, 2, \cdots, s-2)$, meets one and only one of the curves $\bar{L}_j$, $j > i$, and that if $(\bar{L}_i \cdot \bar{L}_{i+\alpha}) = 1$, $\alpha > 0$, then $(\bar{L}_i^2) = -\alpha - 1$. Finally if $(L_i \cdot L_{i+\alpha}) = 1$, then by (ii) $(L_j \cdot L_{j+1}) = 1$ $(j = i+1, \cdots, i+\alpha-2)$, and consequently $(\bar{L}_j \cdot \bar{L}_{j+1}) = 1$ for the same values of j. Hence all the conditions (i), (ii), (iii) are satisfied by the ordered sequence $\bar{L}_1, \cdots, \bar{L}_{s-1}$.

From the hypothesis that our theorem is true for $s-1$, it follows that a proper combination $\bar{L} = \bar{k}_1\bar{L}_1 + \cdots + \bar{k}_{s-1}\bar{L}_{s-1}$ of the curves $\bar{L}_i$ is an exceptional curve of the first kind. There exists then a linear system $\bar{\Sigma}$ on $\bar{F}$ possessing $\bar{L}$ as a total fundamental curve, satisfying the conditions (i), (ii), and (iii) of the definition given in section 1 and *having no base points on* $\bar{L}$. It is seen immediately that the curves $L_1, \cdots, L_s$ are fundamental curves of the system Σ on F which corresponds to $\bar{\Sigma}$, that a proper combination $k_1L_1 + \cdots + k_sL_s$ of these curves is a total fundamental curve of Σ (since by the conditions (i), (ii), (iii) of the preceding theorem the curve $L_1 + \cdots + L_s$ is connected) and that Σ satisfies the conditions (i), (ii), (iii) of the definition given in section 1. It has been proved already (section 7, Remark 2) that the integers k_i are uniquely determined. It remains to prove that the arrangement $L_1, L_2, \cdots, L_s$ is uniquely determined. Let us assume that $L_{i_1}, L_{i_2}, \cdots, L_{i_s}$ is another arrangement satisfying the conditions (i), (ii) and (iii), where $i_1, i_2, \cdots, i_s$ is a permutation of the indices $1, 2, \cdots, s$. Since by (iii) $(L_{i_s}^2) = -1$ and since by (ii) L_s is the only curve whose virtual degree is -1, it follows that L_{i_s} coincides with L_s. We shall therefore assume that $i_j = j$ for $j = \alpha, \alpha+1, \cdots, s$ $(1 < \alpha \leqq s)$, and we shall prove that $i_{\alpha-1} = \alpha - 1$. Let $i_\beta = \alpha - 1$, and let $(L_{\alpha-1} \cdot L_{\alpha+\gamma}) = 1$, $\gamma \geqq 0$, by (ii). We have, by (ii), $(L^2_{\alpha-1}) = -(2+\gamma)$, and by the same condition (ii), applied to the arrangement $L_{i_1}, L_{i_2}, \cdots, L_{i_{\alpha-1}}, L_\alpha, L_{\alpha+1}, \cdots, L_s$, we have $(L^2_{\alpha-1}) = (L^2_{i_\beta}) = -(\alpha+\gamma-\beta+1)$. Hence $2+\gamma = \alpha+\gamma-\beta+1$, or $\beta = \alpha - 1$, q. e. d.

10. *Fundamental points on several irreducible branches.* In the previous sections we have treated the case of a birational transformation with a succession of infinitely near fundamental points on a single irreducible algebroid branch. We found it necessary to assume that the curves $L_i + \cdots + L_s$ $(i = 1, \cdots, s)$ were connected. When these curves were not connected we saw in section 3 that the aggregate of fundamental points would not lie all on one branch. It is the purpose of this section to consider this latter case.

The curve $L_1 + \cdots + L_s$ is connected (see 2), but we shall suppose that $L_2 + \cdots + L_s$ breaks up into ρ (≥ 1) maximal connected components $\bar{L}^{(1)}, \cdots, \bar{L}^{(\rho)}$, having two by two no points in common (see 3). Then on the surface $\bar{F}$, in the first neighborhood of the point O, we have the ρ distinct fundamental points $O_{1,1}, \cdots, O_{\rho,1}$ of the transformation T^{-1}. We shall fix our attention on one of these, say $O_{1,1}$. In the immediate neighborhood of $O_{1,1}$ T^{-1} may have several fundamental points, among which there may or may not be *the* * proximate point of O following $O_{1,1}$. We shall suppose that T^{-1} has a fundamental point $O_{1,2}$, immediately following $O_{1,1}$, which is a proximate point of O. Then in the immediate neighborhood of $O_{1,2}$, T^{-1} may have several fundamental points, among which there may or may not be the proximate point of O following $O_{1,2}$. We shall suppose that T^{-1} has a fundamental point $O_{1,3}$ immediately following $O_{1,2}$, which is a proximate point of O. We continue in this way to define the sequence of points $O, O_{1,1}, \cdots, O_{1,\sigma}$ such that each point is fundamental for T^{-1} and that $O_{1,i}$ is the proximate point of O immediately following $O_{1,i-1}$ $(i = 2, \cdots, \sigma)$.

We shall suppose that the fundamental points of T^{-1} which follow $O_{1,\sigma}$ include no proximate point of O. From the preceding sections it follows that the immediate neighborhood of each point of the sequence $O, O_{1,1}, \cdots, O_{1,\sigma}$ is represented in $(1, 1)$ fashion by the points of irreducible components of L, namely $L_1, L_{1,1}, \cdots, L_{1,\sigma}$ respectively. Consider now a linear system of curves Φ on $\bar{F}$, cut out by a system of forms of a sufficiently high order, having base points only at the points $O, O_{1,1}, \cdots, O_{1,\sigma}$ and free from fundamental curves on $\bar{F}$. By relating the system Φ to the hyperplanes of a linear space, we induce a transformation T' and transform $\bar{F}$ into a surface F'; from the considerations in sections 1-9, it is clear that F' has an exceptional curve of the type treated in those sections. F' is the transform of F by the product

* We say "*the* proximate point of O," because it can be easily shown that there is only one such point in the immediate neighborhood of $O_{1,1}$. More generally, it is not difficult to prove that if $O_i, O_{i+1}, \cdots, O_{i+\beta}$ $(\beta > 0)$ is a sequence of successive points in the neighborhoods of order $i, i+1, \cdots, i+\beta$ of a point O, such that $O_{i+\beta}$ is a proximate point of O_i, then there is one and only one point $O_{i+\beta+1}$ immediately following $O_{i+\beta}$, which is a proximate point of O_i.

$T \cdot T'$, and this transformation is free from fundamental points on F. The exceptional curve on F' consists of the curves $L'_1, L'_{1,1}, \cdots, L'_{1,\sigma}$, the transforms of $L_1, \cdots, L_{1,\sigma}$ on F; the remaining components of L on F are replaced by points of F'.

Now in the exceptional curve on F', L'_1 must intersect $L'_{1,\sigma}$ by Theorem 5, since on $\bar{F}$ $O_{1,\sigma}$ is the last proximate point of O. Does this imply that the corresponding curves L_1 and $L_{1,\sigma}$ on F intersect? Let us assume that the intersection of L'_1 and $L'_{1,\sigma}$ arises from the intersection of L_1 and $L_{1,\sigma}$ on F with a curve L_β which has been replaced by a point on F' and let O_β be the fundamental point on $\bar{F}$ whose immediate neighborhood is in $(1,1)$ correspondence with the points of L_β. This point is not in the sequence $O, O_{1,1}, \cdots, O_{1,\sigma}$, since O_β is transformed by T' into a point of F'. On the other hand the fact that $L_{1,\sigma}$ and L_β intersect implies that if $L^{(j)}$ $(j = 0, 1, \cdots, \sigma - 1)$ denotes the exceptional curve on F which corresponds to the point $O_{1,j}$ $(O_{1,j} \equiv O,\ L^{(0)} \equiv L)$, then $L_{1,\sigma}$ and L_β belong to one and the same maximal connected component of $L^{(j)} - L_{1,j}$, i. e. to $L^{(j+1)}$. Hence there exists a branch passing through $O, \cdots, O_{1,\sigma}, O_\gamma, \cdots, O_\beta$. Set up a transformation T'', as T' was determined, possessing only $O, \cdots, O_\beta$ as fundamental points and transform $\bar{F}$ into F''; F'' has an exceptional curve of the type considered in sections 4-9. In this exceptional curve denote by L_j'' the transform of L_j on F. Then $(L''_1 \cdot L''_{1,\sigma}) = 1$ by Theorem 5 and $(L_1'' \cdot L''_\beta) = 1$ since $(L_1 \cdot L_\beta) = 1$ on F and since the transformation TT'' is free from fundamental points on F. But this is impossible, since L_1'' cannot meet two distinct irreducible components of the exceptional curve on F''.

Hence we must have on F, $(L_1 \cdot L_{1,\sigma}) = 1$. This constitutes an extension of Theorem 5 in whatever way the maximal connected components of $L_i + \cdots + L_s$ $(i = 2, \cdots, s)$ may decompose. In all cases we have shown that L_1 *intersects that curve which is the map of the immediate neighborhood of the fundamental point, which is the last proximate point of* O.

In Theorem 5 we were also able to derive the virtual degree of L_1 and hence the virtual degree of any component of L; we found $(L_1^2) = -(1 +$ number of proximate points of $O)$; we wish to prove this for the present case when the fundamental points lie on several irreducible branches.*

The curve $\bar{L} = L_1 + \cdots + L_s$ is connected, but $L_2 + \cdots + L_s$ breaks up into $\rho \geq 1$ maximal connected components $\bar{L}^{(1)}, \cdots, \bar{L}^{(\rho)}$; L_1 meets $\bar{L}^{(\beta)}$ $(\beta = 1, \cdots, \rho)$ in one point each, and each $\bar{L}^{(\beta)}$ gives rise to an exceptional curve $L^{(\beta)}$ and to a fundamental point $O_{\beta,1}$ in the first neighborhood of O.

* The set of proximate points of a given point O, when several branches pass through O, is by definition the set of all the proximate points of O on the different branches through O ([4], Chap. I, 2).

We shall fix our attention on one of these points, say $O_{1,1}$, and on the corresponding curve $\bar{L}^{(1)} = L_2 + \cdots + L_s$. The intersection $(L_1 \cdot \bar{L}^{(1)}) = 1$ arises from the intersection of L_1 with some curve L_{a_α} of $\bar{L}^{(1)}$ such that $(L_1 \cdot L_{a_\alpha}) = 1$. Now $\bar{L}^{(1)} - L_2$ decomposes into several maximal connected components, one of which contains L_{a_α}; let this component be $\bar{L}^{(2)}_{a_2} = L_{a_2} + \cdots + L_{a_\alpha} + \cdots$; we have $(L_2 \cdot \bar{L}^{(2)}_{a_2}) = 1$ [as $(L_1 \cdot \bar{L}^{(1)}) = 1$] and $L^{(2)}_{a_2}$ gives rise to a fundamental point $O_{1,2}$ in the immediate neighborhood of $O_{1,1}$. Then $\bar{L}^{(2)}_{a_2} - L_{a_2}$ breaks up into several maximal connected components, one of which contains L_{a_α}; let this component be $\bar{L}^{(3)}_{a_3} = L_{a_3} + \cdots + L_{a_\alpha} + \cdots$. As before, there is a fundamental point $O_{1,3}$ in the immediate neighborhood of $O_{1,2}$. We continue in this manner and come finally to a fundamental curve $\bar{L}^{(\alpha)}_{a_\alpha} = L_{a_\alpha} + \cdots$ and this gives rise to a fundamental point $O_{1,\alpha}$ in the neighborhood of order α of O. From the proof in the earlier part of this section it follows that $O_{1,\alpha}$ is the last proximate point of O after $O_{1,1}$ among the fundamental points of T^{-1}.

The proof of section 5 extends directly to show that the immediate neighborhood of $O_{1,\alpha}$ is in $(1,1)$ correspondence with the points of the curve L_{a_α}; then the proof of section 6 follows and shows that the branch of minimum order through $O, O_{1,1}, \cdots, O_{1,\alpha}$ has at these points the multiplicities $k_{a_\alpha}, k^{(1)}_{a_\alpha}, \cdots, k^{(\alpha)}_{a_\alpha} = 1$, where these values are the coefficients of L_{a_α} in the exceptional curves $L, L^{(1)}, L^{(2)}_{a_2}, \cdots, L^{(\alpha)}_{a_\alpha}$, derived from the fundamental curves $\bar{L}, \bar{L}^{(1)}, \cdots, \bar{L}^{(\alpha)}_{a_\alpha}$. But $k^{(1)}_{a_\alpha} = 1$ by (10) of section 3. Therefore $k^{(2)}_{a_\alpha} = \cdots = k^{(\alpha)}_{a_\alpha} = 1$. Since the multiplicity of a point is equal to the sum of the multiplicities of its proximate points, we have at once $k_{a_\alpha} = \alpha$. We apply this procedure to the curves $L^{(\beta)}$ $(\beta = 1, \cdots, \rho)$, and from the relation $(L \cdot L_1) = -1$ we derive in all cases the desired result: $(L_1^2) = -(1 + \text{number of proximate points of } O)$. As in Theorem 5, the proof extends to the other irreducible components of L.

REFERENCES.

[1] G. Castelnuovo and F. Enriques, "Sopra alcune questioni fondamentali nella teoria delle superficie algebriche," *Annali di Matematica pura ed applicata*, Ser. 3, vol. 6 (1901), pp. 165-225.

[2] F. Enriques, "Sulla proprietà caratteristica delle superficie algebriche irregolari," *Rendiconti della Reale Accademia delle scienze di Bologna*, nuova serie, vol. 9 (1904-05).

[3] F. Enriques and O. Chisini, *Lezioni sulla teoria geometrica delle equazioni e delle funzioni algebriche*, vol. 2.

[4] O. Zariski, *Algebraic Surfaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. III, 5, Berlin, Springer (1934).

THE JOHNS HOPKINS UNIVERSITY.

THE TOPOLOGICAL TRANSFORMATIONS OF A SIMPLE CLOSED CURVE INTO ITSELF.

By E. R. VAN KAMPEN.

I. The theory of a topological transformation of a simple closed curve into itself has been developed first of all by Poincaré (a), who needed it for his investigations on the differential equations on a torus. Later P. Bohl (b) formulated a result closely related to the sufficiency in our (13). Then H. Kneser (c) and J. Nielsen (d) gave improved proofs of the existence of the number τ (3) and of the transformation functions $y = g(x)$ and $x = h(y)$ (V). In this paper we have used Nielsen's proof that was published in the Danish language. Denjoy (e, f) found his sufficient condition for the transitive case (15) using continued fractions. Kneser (g) published a theorem equivalent with the necessity in (13). Wintner and Lewis (h) laid the foundation for (14). The papers a, b, f, g consider the question as a problem for differential equations on the torus.

In this paper we give the theory for a simple closed curve only, considerably simplified and completed on several points. The consequences for the solutions of a differential equation on the torus are immediate. The examples that have been assembled will not be repeated here. An account will be found in Denjoy's paper.

II. We consider a simple closed curve C and on it a coördinate x taking all real values, such that two values of x correspond to the same point of C if and only if their difference is an integer.

If a topological transformation of C into itself changes the orientation of C then the transformation has exactly two fixpoints and interchanges the two arcs determined by the fixpoints on C. We exclude this case completely.

Any other topological transformation T of C into itself can be determined by a monotonically increasing continuous real function $f(x)$, satisfying the relation $f(x+1) = f(x) + 1$. However if T is known the function $f(x)$ is not uniquely determined but can be replaced by $f(x) + n$, where n is an arbitrary integer. We can consider $f(x)$ as a topological transformation of the line of real numbers into itself.

If we denote by T^n and f^n the n-th powers of the transformations T and f, then we can represent T^n by any of the functions $f^n(x) + k$, where k is an arbitrary integer.

If p is any point of C and x any real number we will sometimes write p_n for $T^n(p)$ and x_n for $f^n(x)$. As T^n is again a topological transformation of C into itself, $f^n(x)$ will be the same type of real function as $f(x)$.

III. In this section we introduce a real number τ characterizing in a very definite topologically invariant sense the average size of the transformation affected by T on the points of C.

(1) *The limit* $\operatorname{Lim}_{|n|\to\infty} (1/n)x_n = \tau$ *exists and is independent of* x.

As soon as we know that the limit exists for a particular value of x, then the rest of the theorem is very simple. For

$$|x_n - y_n| < |x - y| + 1, \quad \text{so} \quad \operatorname{Lim}_{|n|\to\infty} (1/n)\cdot(x_n - y_n) = 0.$$

If some power T^m of T has a fixpoint with the coördinate y, then we can find an integer p such that $y_m = y + p$. Using the formula $f^m(x+1) = f^m(x) + 1$ we can prove by induction $y_{mn} = y + pn$ and a simple limiting process shows that in this case τ exists and is equal to the rational number p/m.

If no power of T has a fixpoint then *no integers* m *and* p *and real number* x *exist such that* $x_m = x + p$. Thus we can find corresponding to any integer m an integer p such that $p + x < x_m < p + x + 1$, for all real x. In fact we only have to find the value of p for one particular value of x and then the inequalities are true for all x because the corresponding equalities never take place. Applying the inequality for $x = 0, 0_m, 0_{2m}, \cdots, 0_{m(n-1)}$ and adding we find:

$$np < 0_{mn} < n(p+1).$$

Comparing this with the original inequality for $x = 0$ we can prove:

$$\left| \frac{1}{mn} 0_{mn} - \frac{1}{n} 0_n \right| < \frac{1}{|m|}.$$

It is clear how this reasoning has to be rearranged if n is negative. If we interchange m and n in the last formula we can conclude:

$$\left| \frac{1}{n} 0_n - \frac{1}{m} 0_m \right| < \frac{1}{|m|} + \frac{1}{|n|},$$

from which the existence of τ follows immediately.

(2) τ *is irrational if and only if no power of* T *has a fixpoint.*

If τ is the invariant of $f(x)$, then $\tau m + k$ is the invariant of $f^m(x) + k$.

So if τ is rational we can find a power $S = T^m$ of T and a corresponding function $g(x) = f^m(x) + k$ such that $\lim_{|n|\to\infty} (1/n) g^n(0) = m\tau + k = 0$. Supposing at the same time that S has no fixpoints, we can assume for reasons of symmetry that $g(x) > x$ for all x. This makes $g_n(0)$ an increasing function of n. But we must have $g_n(0) < 1$, otherwise we could conclude $m\tau + k > 1/n$. So the sequence $g_n(0)$ has a cluster point x and x must necessarily determine a fixpoint of S contrary to our supposition. So τ must be irrational if no power of T has a fixpoint. From now on we will suppose that τ is irrational.

(3) *The number systems $n\tau + m$ and $x_n + m$ are in a one-to-one monotonous correspondence.*

Here x is arbitrary fixed, m and n take all integer values. We note first that the particular value of x chosen has no influence on the order of the elements in $x_n + m$. This depends on the fact that $x_n - x_k$ can never take an integer value. Choosing $x = 0$ we have to prove that $0_n + m < 0_k + l$ and $n\tau + m < k\tau + l$ are equivalent inequalities. The first can be brought into the form $0_{n-k} < l - m$. As in the proof of (1) this leads to $\tau(n - k) < l - m$. And the conclusion is reversible because $0_{n-k} > l - m$ would lead to $\tau(n - k) > l - m$.

(4) *The residual class of τ modulo one is a topological invariant of T and the oriented curve C.*

We know that the systems of numbers $n\tau + m$ and $0_n + m$ are ordered the same way for each coördinate system of C. If we normalize the functions of x belonging to different coördinate systems with the same orientation by the condition $0 < 0_1 < 1$ then the order of the system $0_n + m$ does not depend on the coördinate system chosen. But then the same is the case for the system $n\tau + m$, so the residual class of τ modulo one does not depend on the coördinate system.

IV. We now consider the set $P = \{p_n\}$ formed by all transforms p_n of p and the set E of all cluster points of P.

(5) *E is invariant under the transformation T and independent of p.*

The first part is obvious because P is invariant under T. The second part follows very easily from the following statement:

(6) *The two closed segments e and e^1 determined on C by two points p_m and*

p_n of P both contain at least one of the transforms q_l of any point q of C. To prove this for e we consider the finite sums $e + T^{m-n}e + T^{2(m-n)}e + \cdots + T^{k(m-n)}e$. As τ is irrational, so different from zero, these sums of adjacent intervals must fill C completely if k is sufficiently large. So one of the intervals $T^{l(m-n)}e$ must contain q and e must contain $q_{l(n-m)}$.

If now the sequence p_{a_ν} converges to r, then there must be a sequence of transforms q_{b_ν} of q converging to the same point r, so E is independent of p.

(7) *E is perfect and either equal to C* (transitive case) *or nowhere dense on C* (intransitive case).

As E is closed the first part follows from (6) if we take q to be a point of E. The second part follows from the fact that P is on C either everywhere dense or nowhere dense. If P is dense in an arbitrary interval of C then we can assume that interval to be e of (6). As a finite number of transforms of that interval fills C completely P must be everywhere dense on C.

V. We now make a closer study of the sets of numbers $A = \{n\tau + m\}$ and $B = \{x'_n + m\}$ mentioned in (3). We take for x' in $x'_n + m$ a fixed but arbitrary number corresponding to one of the points of E.

We are going to define a real function $y = g(x)$. For any number in B we put the function equal to the corresponding number in A. Because of (3) and because A is dense on the real axis the function is continuous and increasing on B. If B is also dense on the real axis (transitive case) this function on B can be extended in a unique way to a continuous increasing function $y = g(x)$, defined for all x. If B is not dense on the real axis (intransitive case) then the function on B can be extended in a unique way to a continuous non-decreasing function $y = g(x)$, constant only on the closed intervals corresponding to the segments of $C - E$.

In the transitive case the function $y = g(x)$ has a continuous increasing inverse $x = h(y)$ transforming the y-axis into the x-axis. In the intransitive case we can define the inverse $x = h(y)$ in such a way that it is increasing and has a countable number of discontinuities corresponding to the countable number of intervals on which $g(x)$ is constant.

The function $y = g(x)$ must satisfy the relations

$$g(x + 1) = g(x) + 1, \tag{8}$$

$$gf(x) = g(x) + \tau, \tag{9}$$

for if x is any number in B and y the corresponding number in A then $y + 1$ corresponds to $x + 1$ and $y + \tau$ corresponds to $f(x) = x_1$.

VI. We can choose the variable y, occurring in $y = g(x)$, as a coördinate proportional to the length of arc on some circle D in a euclidean plane. Then (8) shows that $y = g(x)$ defines a univalued and continuous transformation S of C into D, such that S transforms E into D and eventual segments of $C - E$ into points of D. If no such segments exist (transitive case) S must be T a topological transformation.

Now (8) and (9) show that the transformation T of C is transformed by S into a rotation of D over the angle $2\pi\tau$, so finally clearing up completely the meaning of the number τ. If the transformation T is transitive T is topologically equivalent with a rotation of a circle over an angle $2\pi\tau$. If T is intransitive the same equivalence is true after an invariant countable set of closed segments on C has been transformed into a countable set of points by a preliminary transformation of C.

VII. The problem now arises to find conditions for a transformation T of C (of which no power has a fixpoint) to be transitive or intransitive. The N. S. condition for T to be transitive is that the transforms p_n of a point p of C are everywhere dense on C (7). The following condition on the same points is more suitable for applications:

(10) *N. and S. for T to be transitive is for the transforms p_n of a fixed point p to have the following property: Whenever the sequence p_{a_ν}, a_ν integers, $\nu = 1, 2, \cdots$, is convergent then the sequences $p_{a_\nu+n}$ depending on n converge uniformly in n.*

The necessity is very simple: If T is transitive the curve C can be obtained as the result of the (uniformly) continuous transformation S^{-1} of D into C (see VI). But if a sequence of numbers τa_ν modulo one converges then $\tau(a_\nu + n)$ modulo one converges uniformly in n, and this uniform convergence is not disturbed by the transformation S^{-1}.

We now prove the sufficiency. If p_{a_ν} converges to an arbitrary point q of E (5), then $\lim_{\nu\to\infty} p_{a_\nu+b_\mu+n} = q_{b_\mu+n}$. If we select the sequence b_μ, such that q_{b_μ} converges then $\lim_{\mu\to\infty}\lim_{\nu\to\infty} p_{a_\nu+b_\mu+n}$ exists. As a consequence of the condition in (10) $p_{a_\nu+b_\mu+n}$ will converge uniformly in n and b_μ so $q_{b_\mu+n}$ will converge uniformly in n. In other words the condition in (10) is satisfied for any point of E if it is satisfied for p. We apply it to the endpoints q and r of an eventual complementary segment e of E in C. If q_{a_ν} converges then r_{a_ν} must converge to the same point (at least if all a_ν are distinct). But q_{n+a_ν} and r_{n+a_ν} could not possibly converge uniformly at the same time to the same point. In fact if we substitute $n = -a_\nu$ we find that for any a_ν there exists an n such that the points q_{n+a_ν} and r_{n+a_ν} are as far apart as the distinct fixed

points q and r. In other words a complementary interval e cannot exist and $C = E$. At the same time we find:

(11) *If T is transitive and for fixed p and sequence of integers a_ν the sequence of points p_{a_ν} converges then all sequences q_{a_ν} converge uniformly in q.*

Remark. The criterion in (10) is in a certain degenerate sense still valid in case τ is rational and even when T does change the orientation of C, if we consider as transitive any transformation for which $E = C$ if τ is irrational and for which some power of T is the identical transformation in the other cases.

VIII. If T is transitive it is topologically equivalent with a rotation. So the next question that arises is to find the equations of the continuous group of transformations of which such a transformation T is one element.

We take a fixed point p on C and two variable points q and r. We will define a point s depending on q and r ($s = f(q, r)$) such that the transformation of that continuous group transforming p into q, transforms r into s.

(12) *If p_{a_ν} converges to q and p_{b_μ} converges to r then $p_{a_\nu + b_\mu}$ converges to $s = f(q, r)$.*

The existence and uniqueness of $s = f(q, r)$ follows from (10). That $s = f(q, r)$ is the equation of the desired continuous group is evident if we apply the transformation S (VI). We see that if Sq and Sr are the results of a rotation of Sp over angles $2\pi\lambda$ and $2\pi\mu$, then $Sf(q, r)$ is the result of a rotation of Sp over an angle $2\pi(\lambda + \mu)$. In particular it follows that $s = f(q, r)$ depends continuously on the two points q and r together. $s = f(q, r)$ satisfies the following relations:

$$s = f(q, r) = f(r, q);\ f_n(q, r) = f(q_n, r);$$
$$f(p, r) = r;\ f(q, f(r, s)) = f(f(q, r), s).$$

All these properties of $f(q, r)$ could be derived directly from (10) and (11).

IX. Condition (10) can be modified if we suppose that the curve C is in the complex plane, so that any point of C is a complex number. But for a slight change, that can easily be accounted for, (10) then says:

(13) *N. and S. for T to be transitive is that the transforms p_n of any point p of C form an almost periodic function on the group of addition of all integers n.**

* See J. von Neumann, "Almost periodic functions in a group I," *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 445-492, and A. Walther, "Fastperiodische Folgen," *Hamburg Abhandlungen*, vol. 6 (1928), pp. 217-234.

The considerations of section VI can then be amplified in the following way: The circle D can be considered as representing the group of rotations round its center. It contains a subgroup simply isomorphic with the group of addition of integers: the group of rotations over angles $2\pi\tau n$. If we let p_n correspond to the rotation $2\pi\tau n$ then according to VI the resulting function can be extended to a continuous function on the group D, in fact to the transformation S^{-1} transforming D into C. From the theory of almost periodic functions it follows now that the two almost periodic functions p_n on the group of integers and S^{-1} on the group of rotations D have the same Fourier sequence. If $2\pi\alpha$ represents a variable angle of rotation on D then the Fourier sequence of S^{-1} must have the form $\sum_{-\infty}^{\infty} a_k \exp(2\pi i\alpha k)$. As we find the values of α corresponding to p_n by putting $\alpha = n\tau$ the Fourier sequence of p_n *will have the form* $\sum_{-\infty}^{\infty} a_k \exp(2\pi i n\tau k)$, in other words:

(14) *The Fourier exponents of the almost periodic function* p_n *of* (13) *will be integer multiples of* τ.

X. The equality of the Fourier coefficients of p_n and S^{-1} can be generalized, expressed in terms of the functions $y = g(x)$ and $x = h(y)$ of V and proved in a very simple way.

We need the fact that the smallest positive remainders (τn) of τn modulo one are uniformly everywhere dense between 0 and 1. So if $t(y)$ is a Riemann integrable function defined for $0 \leq y \leq 1$ then $\int_0^1 t(y)\,dy = \operatorname{Lim}_{n\to\infty}(1/n)\sum_{k=1}^{n} t((k\tau))$. If $t(x)$ is everywhere defined, Riemann integrable and has the period one this becomes:

$$\int_0^1 t(y)\,dy = \operatorname*{Lim}_{n\to\infty}\,(1/n)\sum_{l=1}^{n} t(l\tau).$$

Writing $t(g(x)) = s(x)$ we find that $s(x)$ is integrable with respect to $g(x)$ in the Riemann-Stieltjes sense, $s(x+1) = s(x) + 1$ and $t(y) = s(h(y))$. Furthermore for a fixed value x' of x: $t(l\tau) = s(x'_l)$. So finally:

$$(14) \qquad \operatorname*{Lim}_{n\to\infty}\,(1/n)\sum_{l=1}^{n} s(x'_l) = \int_0^1 s(x)\,dg(x) = \int_0^1 s(h(y))\,dy.$$

Under the condition that $s(x+1) = s(x) + 1$ and that either one of the two integrals exists. In the transitive case the equality of the Fourier coefficients follows if we substitute:

$$s(x) = r(x)\exp[2\pi i n g(x)], \qquad s(x'_1) = r(x'_1)\exp[2\pi i n l\tau],$$

where $r(x)$ is the (periodic) function locating C in the complex plane.

XI. The condition (10) does not give a direct connection between the transformation T or the function $f(x)$ determining T, and the transitivity or intransitivity of T. The problem of finding such a condition (at least without further restrictions) seems to be extremely difficult. In the same class are all problems of the type: What can be said about the real function character of $y = g(x)$ and $x = h(y)$ (IV) if something is known about the character of $f(x)$.

The difficulty of this problem is indicated by the following statement of which we omit the simple proof:

If two segments e and e' on C have no point in common and a topological transformation T' of e into e' is given that does not change the direction, then a transformation T of C can be constructed such that $T'p = Tp$ if p is in e, T belongs to an arbitrary τ, and (except when $\tau = 0$) T is either transitive or intransitive.

This points out that, unless strong restrictions are made on the class of functions considered, no condition, that determines the local character of $f(x)$, can have an influence on the transitivity of T or even on the value of τ.

In the next section we will give a proof of the only result available along the lines indicated in this section, the following theorem of Denjoy:

(15) *If the function $f(x)$ determining T has a continuous derivative $f'(x) \neq 0$ and of bounded variation in the segment $0 \leqq x \leqq 1$ then T is transitive.*

XII. We take on C a fixed point p and a segment e with p as endpoint. Then we determine a positive integer n such that either p_n or p_{-n} is the only point p_k, $|k| \leq n$ in the interior of e. We can of course find arbitrarily high such integers n by taking e sufficiently small. We only consider the case that p_{-n} is the point contained in e, then:

(16) *The two finite sequences:*

$$p_0, p_1, \cdots, p_{n-1};$$
$$p_{-n}, p_{1-n}, \cdots, p_{-1},$$

*alternate on the curve C.**

We must show that if $0 \leq k < n$, no point of the two sequences is in the

* The simplified proof of this statement here given was also found by R. H. Fox.

interior of the segment $p_k p_{k-n}$ that has the same orientation as $p_0 p_{-n}$. We consider two cases:

a. p_l is in the segment $p_k p_{k-n}$ and $n > l \geq k - n$. Then $T^{-k}p_l$ is in $T^{-k}(p_k p_{k-n})$ or p_{l-k} in $p_0 p_{-n}$, which has been excluded for $n > l - k \geq -n$.

b. p_l is in $p_k p_{k-n}$ and $-n \leq l < k - n < 0$. Then p_k is in the segment $p_{l+n}p_l$, so p_{k-l-n} is in $p_0 p_{-n}$ and this has again been excluded for $0 < k - l - n < -l \leq n$.

From the assumptions of (15) we can conclude that $F(x) = \log f'(x)$ is continuous, has the period one and is of bounded variation in the segment $0 \leq x \leq 1$. If x is a number corresponding to p we can conclude from (16) that for some constant V:

$$|F(x) + F(x_1) + \cdots + F(x_{n-1}) - (F(x_{-n}) + F(x_{1-n}) + \cdots + F(x_{-1}))| < V.$$

Substituting the definition of $F(x)$ and contracting the logarithms we find:

$$\left| \log \left(\frac{dx_n}{dx} \frac{dx_{-n}}{dx} \right) \right| < V,$$

$$e^{-V} < \frac{dx_n}{dx} \frac{dx_{-n}}{dx} < e^{+V}.$$

Now we take a segment e on C. We denote its length by μ, the length of $T^n e$ by μ_n and an integral over a segment of the real axis corresponding to e by $\int_e dx$. Then:

$$\mu_n = \int \frac{dx_n}{dx}\, dx, \qquad \mu_{-n} = \int_e \frac{dx_{-n}}{dx}\, dx;$$

$$\mu_n + \mu_{-n} = \int_e \left(\frac{dx_n}{dx} + \frac{dx_{-n}}{dx} \right) dx \geq 2 \int_e \left(\frac{dx_n}{dx} \frac{dx_{-n}}{dx} \right)^{1/2} dx > 2\mu e^{-V/2};$$

$$\sum_{-n}^{+n} \mu_k > 2n\mu\, e^{-V/2}.$$

As the sum of the measures of the segments $T^n e$ is not bounded some of these segments must overlap on C and as this is not the case for segments of $C - E$ such segments cannot exist and T must be transitive.

One consequence is that T is transitive if $f(x)$ and its inverse function are analytic.

XIII. The difficulties found in trying to conclude anything about $g(x)$

from the real function character of $f(x)$, disappears completely if we start from the function $s = f(q, r)$.

We suppose that T is transitive, so that the functions $g(x)$ and $h(y)$ are both continuous, and monotonic increasing. Then we define a function $f(x, y)$ of two variables by the equation:

$$(17) \qquad f(x, y) = h[g(x) + g(y)] \text{ or } gf(x, y) = g(x) + g(y).$$

It is quite obvious that $f(x, y)$ is continuous and has the period 1 in both variables. Its meaning on the curve C is given by the function $f(q, r)$ of (12) if we take there for p the point with coördinate zero.

(18) a. *If $f(x, y)$ has the derivative $f_x(x, y)$ then the derivatives $g'(x)$ and $h'(x)$ of $g(x)$ and $h(x)$ exist and are positive.*

b. *If $f_x(x, y)$ is continuous in y then $g'(x)$ and $h'(x)$ are continuous.*

c. *If $f_x(x, y)$ has an n-th derivative to y, then $g^{(n+1)}(x)$ exists, $n > 0$.*

d. *If $f(x, y)$ is analytic then $g(x)$ and $h(x)$ are analytic.*

a. As $g(x)$ is continuous and monotonic increasing it has certainly a derivative for some value z of x. For any given x we can find a value y such that $z = f(x, y)$. For the values of x and y so determined we take the derivative of $gf(x, y) = g(x) + g(y)$ to x and find that $g(x)$ has a derivative for all values of x. Then the formula:

$$(19) \qquad g'f(x, y) \cdot f_x(x, y) = g'(x),$$

now valid for all x and y shows that if $g'(x)$ was zero for one particular value of x it would always be zero and that is obviously impossible. So $g'(x)$ is positive, the inverse function $h(x)$ has a positive derivative and $f(x, y)$ has a positive total derivative.

b. If $f_x(x, y)$ is continuous in y then (19) shows, taking x fixed and $f(x, y) = z$ arbitrary, that $1/g'(x)$ and $h'(z)$ are continuous. As $1/g'(x)$ can never be equal to zero $g'(z)$ is also continuous. Finally we see from (19) that $f_x(x, y)$ has a positive greatest lower bound.

c. If $f_x(x, y)$ has an n-th derivative to y, $n > 0$, then it is continuous in y, so $f_x(x, y) > \epsilon > 0$. Now we can successively take n times the derivative of (19) to y, each time proving that the next derivative of $g(x)$ exists. In fact each time it occurs in exactly one term and with a positive coefficient.

d. If $f(x, y)$ is analytic then $f_y(x, y)$ is continuous and positive. So for

constant $z = f(x, y)$ we can solve for y as an analytic function of x. Then (19) shows that $g'(x)$ is an analytic function of x. The same is then true for $g(x)$ and $h(x)$.

LITERATURE.

a. H. Poincaré, *Oeuvres completes*, t. 1, pp. 137-158.

b. P. Bohl, *Acta Mathematica*, vol. 40 (1916), pp. 321-336.

c. H. Kneser, "Reguläre Kurvenscharen auf den Ringflächen," *Mathematische Annalen*, vol. 91 (1924), pp. 135-154.

d. J. Nielsen, "On topologiske Afbildninger . . . ," *Matematisk Tidsskrift*, B. (1928), pp. 39-46.

e. A. Denjoy, *Comptes Rendus*, vol. 194 (1932), pp. 830-833 and pp. 2014-2016.

f. A. Denjoy, "Sur les courbes définies par les équations differentielles a la surface du tore," *Journal de Mathématique* (9), t. 11 (1932), pp. 333-375.

g. H. Kneser, "Periodische Differentialgleichungen . . . ," *Annali di Mathematica*, vol. 11 (1933), pp. 181-185.

h. D. C. Lewis and A. Wintner, "On the distribution of successive images in the Poincaré transformation problem of a circle into itself," *American Journal of Mathematics*, vol. 56 (1934), pp. 407-410.

THE JOHNS HOPKINS UNIVERSITY.

A THEORY OF POSITIVE INTEGERS IN FORMAL LOGIC.* PART I.

By S. C. KLEENE.

---

1. Introduction. In this paper we shall be concerned primarily with the development of the system of logic based on a set of postulates proposed by A. Church.† Our object is to demonstrate empirically that the system is adequate for the theory of positive integers, by exhibiting a construction of a significant portion of the theory within the system. By carrying out the construction on the basis of a certain subset of Church's formal axioms, we show that this portion at least of the theory of positive integers can be deduced from logic without the use of the notions of *negation*, *class*, and *description*.

Instead of limiting our discussion to the system of Church, we shall employ his rules of procedure in a generalized form, and present our results as valid for a logic based on these rules and any set of well-formed formal axioms which includes 1, 3-11, 14-16.‡ This program is explained in the introductory section of a previous article.§

We presuppose familiarity with the contents of the first article of Church, and of §§ 1-5 of our previous article.

As has been noted, significances can be assigned to the symbols of the logic in such a way that the formulas become assertions of logical truths. The mathematical interest arises from the fact that among the logical entities of the system it is possible to select certain ones which occur in relations of the same form as the relations between certain entities in mathematical theories. Hence if the mathematical entities are identified with, or defined to be, the logical entities, the propositions in which they occur will read as theorems of mathematics. It is in this sense that we are to develop or deduce the theory of positive integers within the logic, *i. e.* we are to define the numbers and

---

* Received October 9, 1933. Revised manuscript received June 18, 1934.

† A. Church, "A set of postulates for the foundation of logic," *Annals of Mathematics*, vol. 33 (1932), pp. 346-366, and a second paper under the same title, vol. 34 (1933), pp. 839-864. We shall refer to these articles by their dates.

‡ Church, 1933, p. 841, or 1932, p. 356.

§ S. C. Kleene, "Proof by cases in formal logic," *Annals of Mathematics*, vol. 35 (1934), no. 3. References to theorems or sections of this paper will be made by prefixing the letter *C* to the number of the theorem or section.

other notions employed in the theory as expressions of the logic, and prove formulas which assert, in the symbolism of the logic, that these expressions stand in the relations which the mathematical theory requires.

2. Equality. The definition

$$= \rightarrow \lambda\mu\nu \cdot \phi(\mu) \supset_{\phi} \phi(\nu),$$

$\{=\}$ (x, y) abbreviated to $[x] = [y]$, and the theorems *

2.1: $\quad {}'x \cdot x = x,$

2.2: $\quad [x = y] \supset_{xy} \cdot y = x,$

2.3: $\quad [x = y] \supset_{xy} \cdot [y = z] \supset_{z} \cdot x = z,$

show that we may carry out within the system certain familiar operations with equalities.

Specifically, it follows from the definition that we may pass from an expression $\boldsymbol{J}$ to an expression $\boldsymbol{J'}$ by the substitution for a free occurrence of $\boldsymbol{N}$ in $\boldsymbol{J}$ of an equal expression $\boldsymbol{N'}$, provided the resulting occurrence of $\boldsymbol{N'}$ in $\boldsymbol{J'}$ is also free, *i. e.* under these circumstances $\boldsymbol{J}, \boldsymbol{N} = \boldsymbol{N'} \vdash \boldsymbol{J'}$. For then, according to C5I, we may convert $\boldsymbol{J}$ into $\{\lambda\boldsymbol{x} \cdot \boldsymbol{K}\}(\boldsymbol{N})$ and $\boldsymbol{J'}$ into $\{\lambda\boldsymbol{x} \cdot \boldsymbol{K}\}(\boldsymbol{N'})$ for a suitably chosen $\boldsymbol{x}$ and $\boldsymbol{K}$, and we can pass from $\{\lambda\boldsymbol{x} \cdot \boldsymbol{K}\}(\boldsymbol{N})$ to $\{\lambda\boldsymbol{x} \cdot \boldsymbol{K}\}(\boldsymbol{N'})$ by means of Rule V using $\boldsymbol{N} = \boldsymbol{N'}$ as major premise. This argument also applies directly to the substitution of occurrences of $\boldsymbol{N'}$ for each of a set of occurrences of $\boldsymbol{N}$, provided the occurrences of $\boldsymbol{N}$ and of $\boldsymbol{N'}$ are free. (Cf. C2IX.) As a special case of this substitution rule we may pass from $\boldsymbol{A} = \boldsymbol{B}$ to $\boldsymbol{A} = \boldsymbol{C}$ where $\boldsymbol{C}$ is obtained by legitimate substitutions of $\boldsymbol{N'}$ for $\boldsymbol{N}$ within $\boldsymbol{B}$.

From 2.1 it follows that we may equate an expression $\boldsymbol{A}$ to itself, or to any expression $\boldsymbol{B}$ obtained from $\boldsymbol{A}$ by conversion, provided only that we can prove $E(\boldsymbol{A})$ or $E(\boldsymbol{B})$. For if we have $E(\boldsymbol{A})$ we infer $\boldsymbol{A} = \boldsymbol{A}$ by 2.1, and convert the $\boldsymbol{A}$ on the right into $\boldsymbol{B}$. If we have $E(\boldsymbol{B})$ we infer $\boldsymbol{B} = \boldsymbol{B}$ by 2.1, and convert the $\boldsymbol{B}$ on the left into $\boldsymbol{A}$. (Cf. § C5, paragraph 2.) C5II characterizes the situations in which we may prove $E(\boldsymbol{M})$.

2.2 enables us to pass from $\boldsymbol{A} = \boldsymbol{B}$ to $\boldsymbol{B} = \boldsymbol{A}$, and 2.3 from $\boldsymbol{A} = \boldsymbol{B}$ and $\boldsymbol{B} = \boldsymbol{C}$ to $\boldsymbol{A} = \boldsymbol{C}$.

The abbreviation $\boldsymbol{A}_1 = \boldsymbol{A}_2 = \boldsymbol{A}_3 = \cdots = \boldsymbol{A}_n$ will be used for $[\boldsymbol{A}_1 = \boldsymbol{A}_2]$

* Church, 1933, Theorems 1, 2, 3.

$[A_2 = A_3]\ [A_3 = A_4] \cdots [A_{n-1} = A_n]$. As a consequence of 2.2 and 2.3, each of the n^2 equalities $A_i = A_j$ $(i, j = 1, \cdots, n)$ is a consequence of the *chain of equalities* $A_1 = A_2 = \cdots = A_n$. Frequently we are interested only in the equality $A_1 = A_n$.

We shall adopt the practice of starting with an expression, A_1, and writing successively $= A_2, = A_3, \cdots, = A_n$, whenever it can be shown that, for an arbitrary F, $F(A_1) \vdash F(A_2)$ and $F(A_2) \vdash F(A_1)$, $F(A_2) \vdash F(A_3)$ and $F(A_3) \vdash F(A_2), \cdots, F(A_{n-1}) \vdash F(A_n)$ and $F(A_n) \vdash F(A_{n-1})$, respectively. Such a *chain of contingent equalities linking* $A_1, A_2, \cdots, A_n$ represents a chain of provable equalities if $E(A_i)$ is provable for any one i $(1 \leq i \leq n)$. For then each of $E(A_1)$, $E(A_2), \cdots, E(A_{n-1})$ follows from $E(A_i)$ and the facts $F(A_1) \vdash F(A_2), \cdots, F(A_n) \vdash F(A_{n-1})$; and then each of the $n - 1$ factors of $A_1 = A_2 = \cdots = A_n$ can be proved with the aid of Theorem I. A special case in which $E(A_i)$ can be proved is that $A_{i-1} = A_i$ or $A_i = A_{i+1}$ is a previously established equality.

The circumstances which we most commonly use to justify linking A_{j+1} to A_j in forming a chain of contingent equalities, *i. e.* circumstances which imply that $F(A_j) \vdash F(A_{j+1})$ and $F(A_{j+1}) \vdash F(A_j)$, are the following: (1) A_j convertible into A_{j+1},* (2) $A_j = A_{j+1}$ or $A_{j+1} = A_j$ previously established, (3) A_{j+1} obtainable from A_j by a substitution of N' for N of the type described above ($N = N'$ or $N' = N$ being previously established).†

Ordinarily we introduce a chain of contingent equalities only as a preliminary to inferring their provability, or their provability from formulas which we have assumed. It is intended that it should be clear from the context when this is the case. It will suffice to point out in the course of building the chain or subsequently that one of the equalities is already known, or that one of the expressions exists. Even this formality may be omitted in situations of a type which has already occurred often, when we evidently make subsequent use of the chain of equalities as proved or as proved under our assumptions.

However chains of contingent equalities are of interest in themselves since a chain of contingent equalities linking N and N' can be used in place of a proved $N = N'$ in passing from J to J' under the conditions above, *i. e.* $J \vdash J'$ can be inferred by means of the former. In particular, in the construction of one chain of contingent equalities we may use, (2′), instead of $A_i = A_{i+1}$ or $A_{i+1} = A_i$ under (2), a previously obtained chain of contingent equalities

* In this case we may write *conv*, instead of $=$, before A_{j+1} in constructing the chain.

† (3) includes (2). The transitive property, $A = B, B = C \vdash A = C$, of equality is a special case of the substitution rule, $J, N = N' \vdash J'$.

linking A_i and A_{i+1}, and, (3′), instead of $N = N'$ or $N' = N$ under (3), a previously obtained chain linking N and N'. Chains of contingent equalities are derived in § 15, under hypotheses which do not imply their provability, with a view to their subsequent use in this manner.

If one chain of contingent equalities is used in obtaining another in the fashion just described, and the latter established as true, then the provability of the former follows. For $E(A_i)$ or $E(N)$ can then be proved according to C5II. Thus starting with conversions and known equalities we can build a hierarchy of contingent equalities. If one of the expressions linked in the chain at the top occurs in a provable formula as a free expression, then all the equalities of the hierarchy are provable.

As with formulas, contingent equalities $M = N$ will often occur below when it is shown merely that they hold as a consequence of "assumptions" $X, Y, Z, \cdots$, *i.e.* when $F(M), X, Y, Z, \cdots \vdash F(N)$ and $F(N), X, Y, Z, \cdots \vdash F(M)$.

When there is ambiguity, we may use an accent to distinguish a contingent equality $M =' N$ from an equality $M = N$. The assertion that $M =' N$ would hold if $X, Y, Z, \cdots$ were added to the list of axioms may be written $X, Y, Z, \cdots \vdash' M =' N$.

3. Definition of the positive integers. Peano's axioms. A construction of the theory of positive integers in the logic has been begun by Church,* who selected formulas to represent the three undefined terms, one, successor, and the notion of being a number, of Peano,† and adopted Peano's definitions of $2, 3, \cdots$ by means of the first two of them, thus:

$$1 \to \lambda f x \cdot f(x).$$
$$S \to \lambda \rho f x \cdot f(\rho(f, x)).$$
$$N \to \lambda \mu \cdot [\phi(1) \cdot \phi(x) \supset_x \phi(S(x))] \supset_\phi \phi(\mu).$$
$$2 \to S(1), \quad 3 \to S(2), \quad 4 \to S(3), \cdots.$$

We follow Church in these definitions, but not in the definition of addition, because J. B. Rosser has proposed one which leads to simpler formal proofs, nor in the definitions of multiplication and subtraction, because it is our program to avoid the use of the function ι.

Our first objective will, of course, be the formulation and proof in the logic of Peano's axioms, except the fourth which involves negation. Formula-

* Church, 1933, § 9.

† G. Peano, *Rivista di Matematica*, vol. 1 (1891), pp. 87-102. Peano called the "first" number "zero."

tions have been given by Church, which we adopt except in the case of the fifth. The fifth we express by 3. 3, which may be a better rendering of Peano's axiom, and which we need in the course of our development of the theory.

The first two, 3. 1 and 3. 2, may be proved thus:

$\mathfrak{A}_1$: $\Sigma(\Sigma)$—by two applications of IV to any axiom.

$\mathfrak{A}_2$: $\{\lambda\rho \cdot \rho(I, \Sigma(\Sigma))\}(1)$—by conversion from $\mathfrak{A}_1$. $(I \rightarrow \lambda x \cdot x.)$

$\mathfrak{A}_3$: $\Sigma y \cdot \{\lambda\rho \cdot \rho(I, \Sigma(\Sigma))\}(y)$—IV, (III, $\mathfrak{A}_2$).

$\mathfrak{D}_y$: $\{\lambda\rho \cdot \rho(I, \Sigma(\Sigma))\}(y)$—assumed, in preparation for an application of Theorem I.

$\mathfrak{C}_{y1}$: $\{\lambda\rho \cdot \rho(I, \Sigma(\Sigma))\}(S(y))$—by conversion from $\mathfrak{D}_y$.

$\mathfrak{A}_4$: $\{\lambda\rho \cdot \rho(I, \Sigma(\Sigma))\}(y) \supset_y \{\lambda\rho \cdot \rho(I, \Sigma(\Sigma))\}(S(y))$—provable, according to Theorem I, using $\mathfrak{A}_3$ and the proof of $\mathfrak{C}_{y1}$ from $\mathfrak{D}_y$.

$\mathfrak{A}_5$: $\Sigma\phi \cdot \phi(1) \cdot \phi(y) \supset_y \phi(S(y))$—IV, III, (14, $\mathfrak{A}_2$, $\mathfrak{A}_4$).

$\mathfrak{D}_\phi$: $\phi(1) \cdot \phi(y) \supset_y \phi(S(y))$—assumed.

$\mathfrak{C}_{\phi 2}$: $\phi(1)$—15, $\mathfrak{D}_\phi$.

$\mathfrak{A}_6$: $[\phi(1) \cdot \phi(y) \supset_y \phi(S(y))] \supset_\phi \phi(1)$—Theorem I, $\mathfrak{A}_5$, $\mathfrak{C}_{\phi 2}$.

3. 1: $N(1)$—III, $\mathfrak{A}_6$.

$\mathfrak{A}_7$: $\Sigma x \cdot N(x)$—IV, (III, 3. 1).

$\mathfrak{D}_x$: $N(x)$—assumed.

$\mathfrak{C}_{\phi x 3}$: $\phi(x)$—V, $\mathfrak{D}_x$, $\mathfrak{D}_\phi$, conversion.

$\mathfrak{C}_{\phi 4}$: $\phi(y) \supset_y \phi(S(y))$—16, $\mathfrak{D}_\phi$.

$\mathfrak{C}_{\phi x 5}$: $\phi(S(x))$—V, $\mathfrak{C}_{\phi 4}$, $\mathfrak{C}_{\phi x 3}$, conversion.

$\mathfrak{C}_{x6}$: $[\phi(1) \cdot \phi(y) \supset_y \phi(S(y))] \supset_\phi \phi(S(x))$—Theorem I, $\mathfrak{A}_5$, $\mathfrak{C}_{\phi x 5}$.

$\mathfrak{C}_{x7}$: $N(S(x))$—III, $\mathfrak{C}_{x6}$.

3. 2: $N(x) \supset_x N(S(x))$—Theorem I, $\mathfrak{A}_7$, $\mathfrak{C}_{x7}$.

The fifth Peano axiom we establish as follows:

$\mathfrak{A}_1$: $N(y)N(y) \supset_y N(S(y))$—provable by Theorem I, using 3. 1, 3. 2, 14, 15.

$\mathfrak{D}_\phi$: $\phi(1) \cdot N(y)\phi(y) \supset_y \phi(S(y))$—assumed; $\Sigma\phi \cdot \mathfrak{D}_\phi$ is provable from 14, 3. 1, $\mathfrak{A}_1$.

$\mathfrak{C}_{\phi 1}$: $N(1)\phi(1)$—14, 3. 1, 15($\mathfrak{D}_\phi$).

$\mathfrak{D}_{\phi,y}$: $N(y)\phi(y)$—assumed; $\Sigma y \cdot \mathfrak{D}_{\phi,y}$ is provable by means of III and IV from $\mathfrak{C}_{\phi 1}$.

$\mathfrak{C}_{\phi y 2}$: $\phi(S(y))$—provable by means of 16($\mathfrak{D}_\phi$) and $\mathfrak{D}_{\phi,y}$.

$\mathfrak{C}_{y3}$: $N(S(y))$—provable by means of 3. 2, 15($\mathfrak{D}_{\phi,y}$).

$\mathfrak{C}_{\phi 4}$: $N(y)\phi(y) \supset_y \cdot N(S(y))\phi(S(y))$—Theorem I, 14($\mathfrak{C}_{\phi y 2}$, $\mathfrak{C}_{y3}$).

$\mathfrak{C}_{\phi 5}$: $N(1)\phi(1) \cdot N(y)\phi(y) \supset_y N(S(y))\phi(S(y))$—14, $\mathfrak{C}_{\phi 1}$, $\mathfrak{C}_{\phi 4}$.

$\mathfrak{D}_x$: $N(x)$—assumed; $\Sigma x \cdot N(x)$ is provable by means of 3. 1.

$\mathfrak{C}_{\phi x 6}$: $N(x)\phi(x)$—V, $\mathfrak{D}_x$, $\mathfrak{C}_{\phi 5}$, conversion.
$\mathfrak{C}_{\phi x 7}$: $\phi(x)$—16, $\mathfrak{C}_{\phi x 6}$.
3.3: $[\phi(1) \cdot N(y)\phi(y) \supset_y \phi(S(y))] \supset_\phi \cdot N(x) \supset_x \phi(x)$ — Theorem I, (Theorem I, $\mathfrak{C}_{\phi x 7}$).

The third Peano axiom will be proved as Theorem 10.1 below.

4. Proof by induction. In the development of the theory of positive integers by means of intuitive logic, the object of Peano's fifth axiom is to justify proofs by induction. By following Frege in making it the definition of x being a positive integer that propositions involving p for which a mathematical induction with respect to p can be carried out should hold when p is taken to be x, we are enabled to express the fifth axiom as a provable formula of the logic. If the logic is adequate, it should then be possible to carry out proofs by induction within the logic. This we show to be the case.

4I. *If* $\boldsymbol{x}$ *is a variable which does not occur in* $\boldsymbol{F}$ *as a free variable, if* $\vdash \boldsymbol{F}(1)$, *and if* $N(\boldsymbol{x}), \boldsymbol{F}(\boldsymbol{x}) \vdash \boldsymbol{F}(S(\boldsymbol{x}))$, *then* $\vdash N(\boldsymbol{x}) \supset_{\boldsymbol{x}} \boldsymbol{F}(\boldsymbol{x})$.

For by means of 14, 3.1, and $\boldsymbol{F}(1)$, we can prove $N(1)\boldsymbol{F}(1)$, and thence, since $\boldsymbol{x}$ does not occur in $\boldsymbol{F}$ as a free symbol and, being a variable, is distinct from the free symbols, Π, &, of N, we can obtain $\{\lambda \boldsymbol{x} \cdot N(\boldsymbol{x})\boldsymbol{F}(\boldsymbol{x})\}(1)$ by conversion. An application of IV gives $\Sigma \boldsymbol{x} \cdot N(\boldsymbol{x})\boldsymbol{F}(\boldsymbol{x})$, and, assuming $N(\boldsymbol{x})\boldsymbol{F}(\boldsymbol{x})$, we can prove $N(\boldsymbol{x})$ and $\boldsymbol{F}(\boldsymbol{x})$ by 15 and 16, and thence, according to the third hypothesis of the Theorem, $\boldsymbol{F}(S(\boldsymbol{x}))$. Then by Theorem I: $N(\boldsymbol{x})\boldsymbol{F}(\boldsymbol{x}) \supset_{\boldsymbol{x}} \boldsymbol{F}(S(\boldsymbol{x}))$. Combining this with $\boldsymbol{F}(1)$ by 14, we obtain a formula convertible into $\{\lambda\phi \cdot \phi(1) \cdot N(y)\phi(y) \supset_y \phi(S(y))\}(\boldsymbol{F})$ since $\boldsymbol{F}, N, \&, S$ do not contain $\boldsymbol{x}$ as a free symbol. Using this formula as minor premise with 3.3 as major premise to an application of Rule V, we obtain $\{\lambda\phi \cdot N(x) \supset_x \phi(x)\}(\boldsymbol{F})$, and by conversion, since $\boldsymbol{F}$ does not contain $\boldsymbol{x}$ as a free variable, $N(\boldsymbol{x}) \supset_{\boldsymbol{x}} \boldsymbol{F}(\boldsymbol{x})$. Under these circumstances we may say that $N(\boldsymbol{x}) \supset_{\boldsymbol{x}} \boldsymbol{F}(\boldsymbol{x})$ is proved by *induction with respect to* $\boldsymbol{x}$ from the *basis* $\boldsymbol{F}(1)$; and we may call $\boldsymbol{F}(\boldsymbol{x})$ the *hypothesis of the induction.* This terminology, with appropriate modifications, will also be used in connection with the generalizations of this theorem.

Certain generalizations of the simple inductive procedure described in 4I are permissible, because they can be reduced to one or more simple inductions. Although the nature of these reductions is quite evident intuitively, we shall give them explicitly to ensure that they can be carried out wholly within the confines of the system.

4II. *If x is a variable which does not occur in $F_1, F_2, \cdots, F_n$ as a free variable, if* $\vdash F_1(1), \vdash F_2(1), \cdots, \vdash F_n(1)$, *and if* $N(x), F_1(x), F_2(x), \cdots, F_n(x) \vdash F_i(S(x))$ $(i = 1, \cdots, n)$, *then* $\vdash N(x) \supset_x F_1(x)$, $\vdash N(x) \supset_x F_2(x), \cdots, \vdash N(x) \supset_x F_n(x)$.

For then we can take $\lambda x \cdot F_1(x) F_2(x) \cdots F_n(x)$ as the F of the simple case, and prove $N(x) \supset_x \cdot F_1(x) F_2(x) \cdots F_n(x)$.* With the aid of this result we can prove each of the theorems $N(x) \supset_x F_1(x), \cdots, N(x) \supset_x F_n(x)$ by means of Theorem I. The restrictions on x have been used tacitly in this argument. A set of theorems which are inferred to be provable by an application of this theorem may be said to be proved by a *simultaneous induction.*

4III. *If x is a variable which does not occur in F as a free variable, if* $\vdash F(1), \vdash F(2), \cdots, \vdash F(n)$, *and if* $N(x), F(x), F(S(x)), \cdots, F(S(\cdots n-1 \text{ times} \cdots (S(x)) \cdots)) \vdash F(S(\cdots n \text{ times} \cdots (S(x)) \cdots))$, *then* $\vdash N(x) \supset_x F(x)$.†

This theorem we establish by an intuitive induction with respect to n. It has been proved as 4I for the case that n is 1. We assume it for a value p of n, and apply it under the hypotheses with n taken as $p+1$ to the function $\lambda x \cdot F(x) F(S(x))$, obtaining $N(x) \supset_x F(x) F(S(x))$. Thence by Theorem I and Axiom 15 we can prove $N(x) \supset_x F(x)$. Under the circumstances of this theorem we may say that $N(x) \supset_x F(x)$ is proved by induction with respect to x from the n-tuple basis $F(1), F(2), \cdots, F(n)$.

4IV. *If $x_1, x_2, \cdots, x_n$ are distinct variables which do not occur as free variables in F, and if* $\vdash F(1, 1, \cdots, 1)$; $N(x_1), F(x_1, 1, \cdots, 1) \vdash F(S(x_1), 1, \cdots, 1)$; $N(x_1), N(x_2), F(x_1, x_2, \cdots, 1) \vdash F(x_1, S(x_2), \cdots, 1)$; $\cdots$; $N(x_1), N(x_2), \cdots, N(x_n), F(x_1, x_2, \cdots, x_n) \vdash F(x_1, x_2, \cdots, S(x_n))$, *then* $\vdash N(x_1) N(x_2) \cdots N(x_n) \supset_{x_1 x_2 \ldots x_n} F(x_1, x_2, \cdots, x_n)$.

This theorem concerning *induction with respect to n variables* reduces to 4I in case n is 1. If the theorem be assumed for the case that n is p, and applied under the hypotheses with n taken as $p+1$ to the function $\lambda x_1 x_2 \cdots x_{n-1} \cdot F(x_1, \cdots, x_{n-1}, 1)$, we obtain $N(x_1) \cdots N(x_{n-1}) \supset_{x_1 \ldots x_{n-1}}$

---

* Henceforth we often omit mention of applications of Axioms 14-16.

† When n is being used to represent a given positive integer of intuitive logic, we may use n to represent the corresponding positive integer $S(\cdots n-1 \text{ times} \cdots (S(1)) \cdots)$ of the formal logic, and *vice versa*. We sometimes employ symbols of the formal theory in a familiar intuitive sense, the context being supposed to indicate when this is being done. For example, 1, n, $+$, $-$ in x_1, x_n, x_{n+1}, $n-1$ times, and n-th.

$\cdot \boldsymbol{F}(\boldsymbol{x}_1 \cdot \cdot \cdot, \boldsymbol{x}_{n-1}, 1)$. Also by means of the last hypothesis and a corollary of Theorem I we obtain that $N(\boldsymbol{x}_n)$, $N(\boldsymbol{x}_1) \cdot \cdot \cdot N(\boldsymbol{x}_{n-1}) \supset_{\boldsymbol{x}_1 \ldots \boldsymbol{x}_{n-1}} \cdot \boldsymbol{F}(\boldsymbol{x}_1, \cdot \cdot \cdot, \boldsymbol{x}_{n-1}, \boldsymbol{x}_n) \vdash N(\boldsymbol{x}_1) \cdot \cdot \cdot N(\boldsymbol{x}_{n-1}) \supset_{\boldsymbol{x}_1 \ldots \boldsymbol{x}_{n-1}} \boldsymbol{F}(\boldsymbol{x}_1, \cdot \cdot \cdot, \boldsymbol{x}_{n-1}, S(\boldsymbol{x}_n))$. Hence, applying 4I to the function $\lambda \boldsymbol{x}_n \cdot N(\boldsymbol{x}_1) \cdot \cdot \cdot N(\boldsymbol{x}_{n-1}) \supset_{\boldsymbol{x}_1 \ldots \boldsymbol{x}_{n-1}} \cdot \boldsymbol{F}(\boldsymbol{x}_1, \cdot \cdot \cdot, \boldsymbol{x}_n)$, we can prove $N(\boldsymbol{x}_n) \supset_{\boldsymbol{x}_n} \cdot N(\boldsymbol{x}_1) \cdot \cdot \cdot N(\boldsymbol{x}_{n-1}) \supset_{\boldsymbol{x}_1 \ldots \boldsymbol{x}_{n-1}} \cdot \boldsymbol{F}(\boldsymbol{x}_1, \cdot \cdot \cdot, \boldsymbol{x}_n)$, and thence, by a corollary of Theorem I, $N(\boldsymbol{x}_1) \cdot \cdot \cdot N(\boldsymbol{x}_n) \supset_{\boldsymbol{x}_1 \ldots \boldsymbol{x}_n} \boldsymbol{F}(\boldsymbol{x}_1, \cdot \cdot \cdot, \boldsymbol{x}_n)$.

Finally we may vary the simple inductive process of 4I in a combination of the three directions represented by 4II, 4III, and 4IV, and justify the procedure by a corresponding combination of the devices used in establishing 4II, 4III, and 4IV.

We shall apply the theorems of this section, and similar theorems below, when the situation exhibited is not exactly that described by the theorem, but one that could be made so by evident conversions. (Cf. § C5.)

5. Addition. We adopt the definition

$$+ \rightarrow \lambda \rho \sigma f x \cdot \rho(f, \sigma(f, x)),$$

due to J. B. Rosser, and abbreviate $\{+\}(\boldsymbol{x}, \boldsymbol{y})$ to $[\boldsymbol{x}] + [\boldsymbol{y}]$.

5I. $\{+\}(1)$ *conv* S. $[\boldsymbol{x} + \boldsymbol{y}] + \boldsymbol{z}$ *conv* $\boldsymbol{x} + [\boldsymbol{y} + \boldsymbol{z}]$. $S(\boldsymbol{x} + \boldsymbol{y})$ *conv* $S(\boldsymbol{x}) + \boldsymbol{y}$.

The last conversion follows from the first two thus: $S(\boldsymbol{x} + \boldsymbol{y})$ conv $1 + \cdot \boldsymbol{x} + \boldsymbol{y}$ conv $[1 + \boldsymbol{x}] + \boldsymbol{y}$ conv $S(\boldsymbol{x}) + \boldsymbol{y}$. It is often more convenient to employ this theorem and like theorems of §§ 6, 7 (often tacitly) than to refer to the formal theorems proved by means of them.

Since $E(S)$ can be proved (*e. g.* from 3. 2), the first conversion of 5I leads, by means of 2. 1,* to

5. 1: $$\{+\}(1) = S.$$

Assume $N(y)$. Then, (1) we can prove $N(S(y))$ by means of 3. 2, and thence, by conversion, $N(1 + y)$, and (2) if we assume $N(x + y)$, we can prove $N(S(x + y))$ by 3. 2, and thence, by conversion, $N(S(x) + y)$. Having (1) and (2), we can prove $N(x) \supset_x N(x + y)$ by induction.† This was done on the assumption $N(y)$, and $\Sigma y \cdot N(y)$ is provable from 3. 1.

---

* Cf. the remarks on 2. 1 in § 2.

† Cf. 4I. In this case only the second of the assumptions $N(\boldsymbol{x})$, $\boldsymbol{F}(\boldsymbol{x})$ is used.

By Theorem I we obtain $N(y) \supset_y \cdot N(x) \supset_x N(x+y)$, and thence, by a corollary of Theorem I,*

5.2: $$N(x)N(y) \supset_{xy} N(x+y).$$

5.2 with 15 and 16 enable us to prove $E([x+y]+z)$ as a consequence of $N(x)N(y)N(z)$, and hence the second conversion of 5I leads to

5.3: $$N(x)N(y)N(z) \supset_{xyz} \cdot [x+y]+z = x+[y+z].$$

We shall use $x+y+z$ as an abbreviation for $[x+y]+z$.

(1) Since $E(2)$ is provable (*e. g.* from 3.1, 3.2) and 2 is convertible into $1+1$, we can prove $1+1=1+1$. (2) Assume $N(x)$ and $x+1=1+x$. Now $S(x)+1$ conv $S(x+1)$, $=S(1+x)$ (by means of the assumption $x+1=1+x$), conv $S(S(x))$ (and $E(S(S(x)))$ is provable from $N(x)$ with the aid of 3.2), conv $1+S(x)$. Hence, by § 2, $S(x)+1=1+S(x)$ is provable from our assumptions. (3) Using (1) and (2) we can prove, by induction

$\mathfrak{A}_1$: $$N(x) \supset_x \cdot x+1=1+x.$$

(4) Assume $N(x)$, $N(y)$, and $x+y=y+x$. Then $x+S(y)$ conv $x+\cdot 1+y$, $=x+\cdot y+1$ (since $y+1=1+y$ is provable from $N(y)$ and $\mathfrak{A}_1$), conv $[x+y]+1$, $=1+[x+y]$ (by means of $\mathfrak{A}_1$, $N(x)$, $N(y)$, 5.2 — this equality is proved from our assumptions), $=1+[y+x]$ (by means of the hypothesis $x+y=y+x$), conv $S(y)+x$. This establishes $x+S(y)=S(y)+x$ from our assumptions, according to § 2. Having (1), (2), and (4), we infer by induction (cf. 4IV)

5.4: $$N(x)N(y) \supset_{xy} \cdot x+y=y+x.$$

Another fundamental theorem on addition will be obtained as 11.4 below.

6. Multiplication. We adopt the definition,

$$\times \rightarrow \lambda\rho\sigma x \cdot \rho(\sigma(x)),$$

of J. B. Rosser, and abbreviate $\{\times\}(x,y)$ to $[x] \times [y]$ or $[x][y]$.†

---

* In making applications of Theorem I and its corollaries we may omit mention of the formulas $\Sigma x \cdot M$, $\Sigma xy \cdot M, \cdots$, when their proof is obvious, as here from 3.1 and 14. (Cf. the statements of Theorem I and its corollaries, Church, 1932, pp. 358, 366.)

† The abbreviation $[x][y]$ for the product of two positive integers, x and y, employed in the presence of $N(x)$ and $N(y)$, should not lead to confusion with the

6I. $[xy]z$ *conv* $x[yz]$. $[x+y]z$ *conv* $xz+yz$.

(1) 3.1 $\vdash E(1)$. $E(1)$, 2.1 $\vdash 1=1$ By conversion from $1=1$, $[1.1]=1$. (2) Assume $N(x)$. Then $E(S(x))$ is provable by means of 3.2. $1S(x)$ conv $S(x)$. Hence $[1S(x)]=S(x)$. From (1) and (2) by induction: *

6.1: $$N(x)\supset_x \cdot 1x = x.$$

Assume $N(y)$. Then (1) by means of 6.1 we have $N(1y)$, and (2) assuming $N(x)$ and $N(xy)$, we can prove $N(1y+xy)$ by means of 5.2, $N(xy)$, and (1), and pass by conversion to $N(S(x)y)$. Having (1) and (2) we infer by induction that $N(y) \vdash N(x)\supset_x N(xy)$. With this result we can prove, by Theorem I Corollary,

6.2: $$N(x)N(y)\supset_{xy} N(xy).$$

Assuming $N(x)N(y)N(z)$, we can prove $E([xy]z)$ by means of 6.2, and $E([x+y]z)$ by means of 5.2 and 6.2. Hence 6I leads to the pair of formal theorems

6.3: $$N(x)N(y)N(z)\supset_{xyz} \cdot [xy]z = x[yz],$$

6.4: $$N(x)N(y)N(z)\supset_{xyz} \cdot [x+y]z = xz+yz.$$

Assume $N(k)$. Then: (1) $1[1+k]=[1\cdot 1]+1k$, by 3.2, 2.1, conversion and 6.1. (2) Assume also $N(l)$ and $l[1+k]=l1+lk$. Then $S(l)[1+k]$ conv $1[1+k]+l[1+k]$, $=1+k+l[1+k]$ (by 6.1, 3.1, 5.2), $=1+k+\cdot l1+lk$ (by our assumption $l[1+k]=l1+lk$), conv $1+[k+l1]+lk$, $=1+[l1+k]+lk$ (by 3.1, 6.2, 5.4), conv $[1\cdot 1]+l1+\cdot k+lk$, $=[1\cdot 1]+l1+\cdot 1k+lk$ (by 6.1), conv $S(l)1+S(l)k$. We can prove the existence of any one of these expressions from our assumptions by means of 3.1, 5.2, 6.2. Hence, by § 2, $S(l)[1+k]=S(l)1+S(l)k$. (3) By induction, using (1) and (2), $N(l)\supset_l \cdot l[1+k]=l1+lk$. According to Theorem I, this argument enables us to infer the provability of

$\mathfrak{A}_1$: $$N(k)\supset_k \cdot N(l)\supset_l \cdot l[1+k]=l1+lk.$$

---

abbreviation for the logical product of two propositions. The abbreviation $[x][y][z]$ for $[[x][y]][z]$ will be employed with the arithmetical product as with the logical product.

* Cf. 4I. In this case only the first of the assumptions $N(x)$ and $F(x)$ is used.

(1) By conversion from $1 = 1 : [1 \cdot 1] = 1 \cdot 1$. (2) Assume $N(x)$ and $x1 = 1x$. Then $S(x)1$ conv $1 + x1$, $= 1 + 1x$ (by the second assumption), $= 1 + x$ (by 6.1 and $N(x)$), conv $S(x)(E(S(x))$ is provable from $N(x)$, using 3.2), conv $1S(x)$. Hence by § 2: $S(x)1 = 1S(x)$. (3) By induction, using (1) and (2),

$\mathfrak{A}_2$: $$N(x) \supset_x \cdot x1 = 1x.$$

(4) Assume $N(x)$, $N(y)$, $xy = yx$. Then $xS(y)$ conv $x[1 + y]$, $= x1 + xy$ (by $\mathfrak{A}_1$, $N(x)$, $N(y)$), $= x1 + yx$ (by the hypothesis $xy = yx$), $= 1x + yx$ (by $\mathfrak{A}_2$, $N(x)$), conv $S(y)x$. Hence $xS(y) = S(y)x$. Having (1), (2), and (4), we can prove, by induction (cf. 4IV),

6.5: $$N(x)N(y) \supset_{xy} \cdot xy = yx.$$

7. Exponentiation. $1(\boldsymbol{F}, \boldsymbol{A})$ conv $\boldsymbol{F}(\boldsymbol{A})$, and $S(\boldsymbol{x}, \boldsymbol{F}, \boldsymbol{A})$ conv $\boldsymbol{F}(\boldsymbol{x}(\boldsymbol{F}, \boldsymbol{A}))$. By an intuitive induction, utilizing the definitions of 2, 3, 4, $\cdots$ from 1, we infer that for any given positive integer z, $\boldsymbol{z}(\boldsymbol{F}, \boldsymbol{A})$ is convertible into $\boldsymbol{F}(\cdots z \text{ times } \cdots(\boldsymbol{F}(\boldsymbol{A}))\cdots)$. Church's definitions of the positive integers were framed with a view to providing this formal means of representing the z-th power of a function $\boldsymbol{F}$ of an argument $\boldsymbol{A}$. We recognize it by introducing the abbreviation $[\boldsymbol{x}]^{[\boldsymbol{p}]}$ for $\{\boldsymbol{p}\}(\boldsymbol{x})$, so that $\boldsymbol{z}(\boldsymbol{F}, \boldsymbol{A})$ may be written $\boldsymbol{F}^{\boldsymbol{z}}(\boldsymbol{A})$.

3.1, 3.2, 14 $\vdash \Sigma\phi af \cdot \phi(a) \cdot \phi(\rho) \supset_\rho \phi(f(\rho))$. Assuming $\phi(a) \cdot \phi(\rho) \supset_\rho \phi(f(\rho))$, we have, by the use of the first factor with the second, $\phi(f(a))$, which is convertible into $\phi(f^1(a))$. Assuming also $\phi(f^x(a))$, the second factor gives $\phi(f(f^x(a)))$, which is convertible into $\phi(f^{S(x)}(a))$. Hence, by induction and the corollary of Theorem I,

7.1: $$[\phi(a) \cdot \phi(\rho) \supset_\rho \phi(f(\rho))] \supset_{\phi af} \cdot N(x) \supset_x \phi(f^x(a)).$$

(1) $I^1 = I$ is provable from $I = I$ by conversion.* (2) Assuming $I^x = I$, we have $I^{S(x)}$ conv $\lambda y \cdot I^x(y)$, $= \lambda y \cdot I(y)$ (by means of our assumption), conv I. $\vdash E(I)$. Hence, by § 2, $I^{S(x)} = I$. From (1) and (2) by induction,

7.2: $$N(x) \supset_x \cdot I^x = I.$$

It follows, since $I(\boldsymbol{A})$ conv $\boldsymbol{A}$, that we can express formally a function of positive integers having any given constant value $\boldsymbol{A}$. Indeed, if

* $I = I$ is provable by 2.1 since $\vdash E(I)$ (cf. §§ C6, C8).

$$\mathfrak{C} \rightarrow \lambda\pi \cdot [\lambda\rho\sigma \cdot I^{\sigma}(\rho)]^{\pi},$$

then $\mathfrak{C}(1, A)$, $\mathfrak{C}(2, A)$, $\cdots$ are functions whose values for any one, two, $\cdots$ members of the sequence $1, 2, \cdots$ (*i. e.* "known" positive integers), respectively, are convertible into A.

The formula 6. 1 is convertible into

7. 3: $$N(x) \supset_x \cdot [\lambda f a \cdot f^x(a)] = x.$$

7I. $F^x(F^y(A))$ *conv* $F^{x+y}(A)$.

Let f and a be proper symbols not occurring in x and y as free symbols. Then the formula $x + y$ is convertible into $\lambda f a \cdot f^x(f^y(a))$ and $\lambda f a \cdot f^{x+y}(a)$. This leads evidently to the conversion 7I, and also shows from a new point of view why the definition given in § 5 for the addition of positive integers is suitable.

The observation is due to J. B. Rosser that if p and q are positive integers the formula x^p is actually the p-th power of x in the arithmetic sense.

7II. $x^p x^q$ *conv* x^{p+q}. $[x^p]^q$ *conv* x^{qp}.

8. Number dyads and triads. A finite ordered set of expressions can be defined intuitively by enumerating its members, $A_1, A_2, \cdots, A_k$. In order to carry out an argument in the logic about such a set whose members vary according to given laws, we may require a formula A such that (1) A is a function of the k members $A_1, \cdots, A_k$, and (2) $A_1, \cdots, A_k$ are k functions of A. If there exist expressions $H_1, \cdots, H_k$ such that $H_1(A_1) = I, \cdots, H_k(A_k) = I$, we can take as A the expression $\lambda x \cdot x(A_1, \cdots, A_k)$, where x is any proper symbol not occurring in $A_1, \cdots, A_k$ as a free symbol. In the important case that $A_1, \cdots, A_k$ are positive integers there are the simpler constructions which follow.

To represent the ordered pair of numbers x and y, we employ the formula $D(x, y)$, abbreviated $[x, y]$, where

$$D \rightarrow \lambda\rho\sigma f g a \cdot f^{\rho}(g^{\sigma}(a)).$$

Then x and y are the functions D_1 and D_2,

$$D_1 \rightarrow \lambda\rho f \cdot \rho(f, I) \quad \text{and} \quad D_2 \rightarrow \lambda\rho \cdot \rho(I),$$

respectively, of $[x, y]$, as is established in the theorems

8.1: $$N(x)N(y) \supset_{xy} \cdot \boldsymbol{D}_1([x, y]) = x,$$

8.2: $$N(x)N(y) \supset_{xy} \cdot \boldsymbol{D}_2([x, y]) = y.$$

The first is provable, according to § 2 and the corollary of Theorem I, since assuming $N(x)N(y)$ we have $\boldsymbol{D}_1([x, y])$ conv $\lambda fa \cdot f^x(I^y(a)), = \lambda fa \cdot f^x(I(a))$ (by 7.2, $N(y)$), conv $\lambda fa \cdot f^x(a)$, $= x$ (by 7.3, $N(x)$); the second, since $\boldsymbol{D}_2([x, y])$ conv $\lambda fa \cdot I^x(f^y(a))$, from which we can pass similarly to y.

Let

$$\boldsymbol{T} \to \lambda\rho\sigma\tau fgha \cdot f^\rho(g^\sigma(h^\tau(a))),$$

and abbreviate $\boldsymbol{T}(x, y, z)$ to $[x, y, z]$. If

$$\boldsymbol{T}_1 \to \lambda\rho f \ \cdot \rho(f, I, I), \quad \boldsymbol{T}_2 \to \lambda\rho f \ \cdot \rho(I, f, I), \quad \boldsymbol{T}_3 \to \lambda\rho \cdot \rho(I, I),$$

then

8.3: $$N(x)N(y)N(z) \supset_{xyz} \cdot \boldsymbol{T}_1([x, y, z]) = x,$$

8.4: $$N(x)N(y)N(z) \supset_{xyz} \cdot \boldsymbol{T}_2([x, y, z]) = y,$$

8.5: $$N(x)N(y)N(z) \supset_{xyz} \cdot \boldsymbol{T}_3([x, y, z]) = z.$$

Similarly for tetrads, etc.

9. Predecessor. Given the third Peano axiom, as formulated in 10.1, a function for the predecessor of a positive integer can be obtained by means of the description, ι. Conversely, given a predecessor function, 10.1 can be proved. By obtaining a predecessor function first, and making it the basis of the proof of 10.1, and of further definitions, we are able to limit our formal assumptions, avoiding, particularly, the use of the description.

The predecessor function, P, which we shall define has, besides the property that *if* $\boldsymbol{x}$ is one of the expressions 1, 2, $\cdots$, then $P(S(\boldsymbol{x}))$ conv $\boldsymbol{x}$, also the property $P(1)$ conv 1. The arithmetical functions, such as subtraction and division, defined by means of it, are likewise given a value within the class of positive integers for every set of positive integral values of the arguments, even in cases when this is not ordinarily done. The resulting special properties of the functions are useful in the development of the theory.

Let

$$\mathfrak{F} \to \lambda\rho \cdot [\boldsymbol{T}_2(\rho), \boldsymbol{T}_3(\rho), S(\boldsymbol{T}_3(\rho))] \quad \text{and} \quad \mathfrak{J} \to [1, 1, 1].$$

Then (1) $\mathfrak{F}^{S(1)}(\mathfrak{J})$ conv $[1, 2, 3]$, and $E([1, 2, 3])$ is provable since

$[1, 2, 3]$ occurs in the proposition $\boldsymbol{T}_1([1, 2, 3]) = 1$ which is provable from 8. 3, 3. 1, and 3. 2. Hence, by 2. 1, $\mathfrak{F}^{S(1)}(\mathfrak{Z}) = [1, S(1), S(S(1))]$. (2) Assume $N(x)$ and $\mathfrak{F}^{S(x)}(\mathfrak{Z}) = [x, S(x), S(S(x))]$. From $N(x)$ we infer $N(S(x))$, $N(S(S(x)))$, and $N(S(S(S(x))))$ by 3. 2. Now $\mathfrak{F}^{S(S(x))}(\mathfrak{Z})$ conv $[\boldsymbol{T}_2(\mathfrak{F}^{S(x)}(\mathfrak{Z})), \boldsymbol{T}_3(\mathfrak{F}^{S(x)}(\mathfrak{Z})), S(\boldsymbol{T}_3(\mathfrak{F}^{S(x)}(\mathfrak{Z})))]$, $= [\boldsymbol{T}_2([x, S(x), S(S(x))]), \boldsymbol{T}_3([x, S(x), S(S(x))]), S(\boldsymbol{T}_3([x, S(x), S(S(x))]))]$ (by the hypothesis of the induction), $= [S(x), S(S(x)), S(S(S(x)))]$ (by means of the equalities obtained by using $N(x)$, $N(S(x))$, and $N(S(S(x)))$ in 8. 4 and 8. 5). $E([S(x), S(S(x)), S(S(S(x)))])$ is given by 8. 3, $N(S(x))$, $N(S(S(x)))$, and $N(S(S(S(x))))$. Hence, by § 2, $\mathfrak{F}^{S(S(x))}(\mathfrak{Z}) = [S(x), S(S(x)), S(S(S(x)))]$. Having (1) and (2), we can prove by induction

$\mathfrak{A}_1$: $$N(x) \supset_x \cdot \mathfrak{F}^{S(x)}(\mathfrak{Z}) = [x, S(x), S(S(x))].$$

Let

$$P \rightarrow \lambda\rho \cdot \boldsymbol{T}_1(\mathfrak{F}^{\rho}(\mathfrak{Z})).$$

Assume $N(x)$. By 3. 2, $N(S(x))$ and $N(S(S(x)))$. Then $P(S(x))$ conv $\boldsymbol{T}_1(\mathfrak{F}^{S(x)}(\mathfrak{Z}))$, $= \boldsymbol{T}_1([x, S(x), S(S(x))])$ (by $N(x)$ and $\mathfrak{A}_1$), $= x$ (by means of 8. 3, $N(x)$, $N(S(x))$, and $N(S(S(x)))$). Hence, by § 2 and Theorem I,

9. 1: $$N(x) \supset_x \cdot P(S(x)) = x.$$

The theorem

9. 2: $$P(1) = 1$$

is provable by conversion from $1 = 1$; and the theorem

9. 3: $$N(x) \supset_x N(P(x))$$

is provable by induction, since (1) $N(P(1))$ follows by conversion from 3. 1, and (2) assuming $N(x)$, we can obtain $P(S(x)) = x$ from 9. 1, and then $N(P(S(x)))$ from $N(x)$.

10. Peano's third axiom. $N(1)N(1) \cdot S(1) = S(1)$ is provable by means of 3. 1, 3. 2, 2. 1, and 14; and thence $\Sigma xy \cdot N(x)N(y) \cdot S(x) = S(y)$ can be proved. Assume $N(x)N(y) \cdot S(x) = S(y)$. Then $x = P(S(x))$ (by 9. 1 and $N(x)$), $= P(S(y))$ (by $S(x) = S(y)$), $= y$ (by 9. 1 and $N(y)$). Hence, by § 2 and Theorem I Cor.,

10. 1: $$[N(x)N(y) \cdot S(x) = S(y)] \supset_{xy} \cdot x = y.$$

11. Subtraction. Let

$$--\rightarrow \lambda\mu\nu \cdot P^{\nu}(\mu),$$

and abbreviate $\{--\}(\boldsymbol{x}, \boldsymbol{y})$ to $[\boldsymbol{x}] - [\boldsymbol{y}]$. If x and y are positive integers, then $\boldsymbol{x} - \boldsymbol{y}$ has the usual significance if x is greater than y, and $\boldsymbol{x} - \boldsymbol{y}$ is 1 if x is less than or equal to y.

$N(x), 9.3, 7.1 \vdash N(y) \supset_y N(x - y)$. Hence, by Theorem I Cor.,

11.1: $$N(x)N(y) \supset_{xy} N(x - y).$$

(1) Assuming $N(x)$, we have $[x + 1] - 1 = [1 + x] - 1$ (by 5.4, $N(x)$, and 3.1), conv $P(S(x))$, $= x$ (by 9.1 and $N(x)$). Hence, by § 2 and Theorem I, $N(x) \supset_x \cdot [x + 1] - 1 = x$. (2) Assume $N(y)$ and $N(x) \supset_x \cdot [x + y] - y = x$. Assume also $N(x)$. Then $[x + S(y)] - S(y)$ $= [S(y) + x] - S(y)$ (by 5.4, $N(y)$, 3.2, $N(x)$), conv $P([1 + y + x] - y)$, $= P([y + 1 + x] - y)$ (by 5.4, $N(y)$, 3.1), conv $P([y + S(x)] - y)$, $= P([S(x) + y] - y)$ (by 5.4, $N(y)$, $N(x)$, 3.2), $= P(S(x))$ (by the second assumption, $N(x)$, 3.2), $= x$ (by 9.1, $N(x)$). Hence, by § 2 and Theorem I, $N(x) \supset_x \cdot [x + S(y)] - S(y) = x$. From (1) and (2) by induction, $N(y) \supset_y \cdot N(x) \supset_x \cdot [x + y] - y = x$; whence, by Theorem I Cor.,

11.2: $$N(x)N(y) \supset_{xy} \cdot [x + y] - y = x.$$

(1) Assume $N(x)N(y)$. Then $[x - y] - 1$ conv $P(y(P, x))$, conv $S(y, P, x)$, conv $x - S(y)$, conv $x - \cdot 1 + y$, $= x - \cdot y + 1$ (by 5.4, 3.1, $N(y)$). $N(x)$, $N(y)$, 3.1, 5.2, 11.1 $\vdash N(x - \cdot y + 1)$, from which $E(x - \cdot y + 1)$ is provable. Hence, by § 2 and Theorem I Cor., $N(x)N(y) \supset_{xy}$ $\cdot [x - y] - 1 = x - \cdot y + 1$. (2) Assume $N(z)$ and $N(x)N(y) \supset_{xy}$ $\cdot [x - y] - z = x - \cdot y + z$. Assume also $N(x)N(y)$. Then $[x - y] - S(z)$ conv $[x - y] - \cdot 1 + z$, $= [[x - y] - 1] - z$ (by the hypothesis of the induction, $N(x), N(y), 11.1, 3.1$), $= [x - \cdot y + 1] - z$ (by (1), $N(x), N(y)$), $= x - \cdot y + 1 + z$ (by the hypothesis of the induction, $N(x)$, $N(y)$, 3.1, 5.2), conv $x - \cdot y + S(z)$. We can prove $E(x - \cdot y + S(z))$ by means of $N(x)$, $N(y)$, $N(z)$, 3.2, 5.2, 11.1. Hence, by § 2 and Theorem I Cor., $N(x)N(y) \supset_{xy} \cdot [x - y] - S(z) = x - \cdot y + S(z)$. Having (1) and (2), we infer by induction $N(z) \supset_z \cdot N(x)N(y) \supset_{xy} \cdot [x - y] - z = x - \cdot y + z$, and thence

11.3: $$N(x)N(y)N(z) \supset_{xyz} \cdot [x - y] - z = x - [y + z].$$

Assume $N(x)N(y)N(z) \cdot x + y = x + z$. Then $y = [x + y] - x$ (by 11.2, 5.4, $N(y)$, $N(x)$), $= [x + z] - x$ (by $x + y = x + z$), $= z$ (by 11.2, 5.4, $N(x)$, $N(z)$). Hence

11.4: $$[N(x)N(y)N(z) \cdot x + y = x + z] \supset_{xyz} \cdot y = z.$$

12. Order. Let $< \rightarrow \lambda\mu\nu \cdot N(\mu) \cdot [\phi(S(\mu)) \cdot \phi(\xi) \supset_\xi \phi(S(\xi))] \supset_\phi \phi(\nu)$, and abbreviate $\{<\}(\boldsymbol{x}, \boldsymbol{y})$ to $\boldsymbol{x} < \boldsymbol{y}$ or $\boldsymbol{y} > \boldsymbol{x}$.

12.1: $$N(x) \supset_x \cdot x < S(x).$$

Proof. Assume $N(x)$. Then, using 3.2, $\Sigma\phi \cdot \phi(S(x)) \cdot \phi(\xi) \supset_\xi \phi(S(\xi)) \cdot \phi(S(x)) \cdot \phi(\xi) \supset_\xi \phi(S(\xi)) \vdash \phi(S(x))$. By Theorem I, $[\phi(S(x)) \cdot \phi(\xi) \supset_\xi \phi(S(\xi))] \supset_\phi \phi(S(x))$, which with $N(x)$ yields $x < S(x)$.*

12.2: $$x < y \supset_{xy} \cdot N(x)N(y).$$

Proof. $12.1 \vdash \Sigma xy \cdot x < y$. $x < y \vdash N(x) \cdot [\phi(S(x)) \cdot \phi(\xi) \supset_\xi \phi(S(\xi))] \supset_\phi \phi(y)$. $N(x)$, 3.2 $\vdash N(S(x))$. $N(S(x))$, 3.2, $[\phi(S(x)) \cdot \phi(\xi) \supset_\xi \phi(S(\xi))] \supset_\phi \phi(y) \vdash N(y)$.

12.3: $$[x < y \cdot y < z] \supset_{xyz} \cdot x < z.$$

Proof. Assume $x < y \cdot y < z$ and $\phi(S(x)) \cdot \phi(\xi) \supset_\xi \phi(S(\xi))$. The latter with $x < y$ yields $\phi(y)$. $\phi(y)$, $\phi(\xi) \supset_\xi \phi(S(\xi)) \vdash \phi(S(y)) \cdot \phi(\xi) \supset_\xi \phi(S(\xi))$, which with $y < z$ yields $\phi(z)$. Hence $[\phi(S(x)) \cdot \phi(\xi) \supset_\xi \phi(S(\xi))] \supset_\phi \phi(z)$, and $x < z$.

When each a_i $(i = 1, \cdots, n)$ is either $=$ or $<$ ($=$ or $>$), we abbreviate $[\boldsymbol{x}_1 a_1 \boldsymbol{x}_2][\boldsymbol{x}_2 a_2 \boldsymbol{x}_3] \cdots [\boldsymbol{x}_n a_n \boldsymbol{x}_{n+1}]$ to $\boldsymbol{x}_1 a_1 \boldsymbol{x}_2 a_2 \boldsymbol{x}_3 \cdots \boldsymbol{x}_n a_n \boldsymbol{x}_{n+1}$. By § 2 and 12.3, we can make formally the usual inferences concerning the order relation of two expressions linked in such a chain of inequalities.†

---

* Henceforth our proofs will be abbreviated by omissions of such details as applications of Theorem I and Corollaries, applications of § 2 and like principles for inequalities, and references to formal theorems under circumstances in which it is clear what theorems are being used. In particular, required formulas of the form $N(\boldsymbol{A})$ will not be mentioned, when they are obtainable from the hypotheses and such theorems as 3.1, 3.2, 5.2, 12.2. In case several theorems are used at a given step in the argument, those playing a subordinate rôle may not be cited.

† We arrange the introduction and proof of chains of inequalities in the same manner as that of chains of equalities. Any link may be a conversion or contingent equality.

12.4: $N(x) \supset_x \cdot N(y) \supset_y \cdot x + y > y.$

Proof. Assume $N(x)$, $N(y) \supset_y \cdot x + y > y$, and $N(y)$. Then $S(x) + y$ conv $S(x + y)$, $> x + y$ (12.1), $> y$ (by means of the hypotheses). Hence, by Theorem I, $N(x)$, $N(y) \supset_y \cdot x + y > y \vdash N(y) \supset_y \cdot S(x) + y > y$. Thus 12.4 is provable by induction from 12.1 as basis.

12.5: $x < y \supset_{xy} \cdot y = [y - x] + x.$

Proof. Assume $x < y$. (1) $S(x)$ conv $1 + x$, $= [[1 + x] - x] + x$ (11.2), conv $[S(x) - x] + x$. Hence $N(S(x)) \cdot S(x) = [S(x) - x] + x$. (2) Assume $N(\xi) \cdot \xi = [\xi - x] + x$. Then $S(\xi) = S([\xi - x] + x)$ (by the hyp.), conv $S(\xi - x) + x$, $= [[S(\xi - x) + x] - x] + x$ (11.2), conv $[S([\xi - x] + x) - x] + x$, $= [S(\xi) - x] + x$ (by the hyp.). Hence $N(S(\xi)) \cdot S(\xi) = [S(\xi) - x] + x$. By Theorem I, $[N(\xi) \cdot \xi = [\xi - x] + x] \supset_\xi \cdot N(S(\xi)) \cdot S(\xi) = [S(\xi) - x] + x$. (3) $x < y \vdash [\phi(S(x)) \cdot \phi(\xi) \supset_\xi \phi(S(\xi))] \supset_\phi \phi(y)$, which with (1) and (2) yields $N(y) \cdot y = [y - x] + x$.

12.6: $x < y \supset_{xy} \cdot x = y - [y - x].$

12.7: $[N(x) \cdot y > z] \supset_{xyz} \cdot [x + y] - z = x + [y - z].$

Proofs. Assuming $x < y$ and letting $\boldsymbol{p} \to y - x$, $x = [\boldsymbol{p} + x] - \boldsymbol{p}$ (11.2), $= [\boldsymbol{p} + x] - \cdot [\boldsymbol{p} + x] - x$ (11.2), $= y - \cdot y - x$ (12.5). Assuming $N(x) \cdot y > z$ and letting $\boldsymbol{p} \to y - z$, $[x + y] - z = [x + \boldsymbol{p} + z] - z$ (12.5), $= x + \boldsymbol{p}$ (11.2), $= x + \cdot [\boldsymbol{p} + z] - z$ (11.2), $= x + \cdot y - z$ (12.5).

12.8: $N(x) \supset_x \cdot S(x) > 1.$

12.9: $N(y) \supset_y \cdot [x > y] \supset_x \cdot x > 1.$

12.10: $N(x) N(y) \supset_{xy} \cdot x + y > 1.$

Proofs. 12.4 $\vdash$ 12.8, and 12.9 is provable by induction, since $[x > 1] \supset_x \cdot x > 1$ and $N(y) \vdash [x > S(y)] \supset_x \cdot x > 1$ (12.4). 12.10 follows from 12.9 by 12.4.

12.11: $[N(x) \cdot y < z] \supset_{xyz} \cdot x + y < x + z.$

12.12: $[N(x) N(y) N(z) \cdot x + y < x + z] \supset_{xyz} \cdot y < z.$

Proofs. Assuming $N(x) \cdot y < z$, then $x + z = x + y + \cdot z - y$ (12.5),

$> x + y$ (12.4). Assuming $N(x)N(y)N(z) \cdot x + y < x + z$, then $z = [x + z] - x$ (11.2), $= [\boldsymbol{p} + x + y] - x$ (by 12.5, if $\boldsymbol{p} \to [x + z] - \cdot x + y$), $= \boldsymbol{p} + \cdot [x + y] - x$ (12.7, 12.4), $= \boldsymbol{p} + y$ (11.2), $> y$ (12.4).

12.13: $$[N(z) \cdot x < S(y) \cdot y < S(z)] \supset_{xyz} \cdot x < S(z).$$

12.14: $$[N(z) \cdot x < y \cdot y < S(z)] \supset_{xyz} \cdot x < z.$$

12.15: $$[x < S(y) \cdot y < z] \supset_{xyz} \cdot x < z.$$

Proofs. Assuming $N(z) \cdot x < S(y) \cdot y < S(z)$, we have $1 + S(z) = 1 + \cdot [S(z) - y] + y$ (12.5), $= [S(z) - y] + S(y)$, $= [S(z) - y] + [S(y) - x] + x$ (12.5), $> 1 + x$ (12.11, 12.10), and hence, by 12.12, $S(z) > x$. Similarly for 12.14 and 12.15.

We let $\boldsymbol{x \leq y} \to \boldsymbol{x < S(y)}$,* and employ the relation $\leq$ as well as $<$ and $=$ in our chains of inequalities. 12.13-12.15 together with previously noted facts show that we can make formally the usual inferences concerning the order relation of two expressions linked in such a chain.

12.16: $$[N(y) \cdot x < S^2(y)] \supset_{xy} \cdot x - y = 1.$$

Proof. Assume $N(y) \cdot x < S^2(y)$. By 11.2, $[S^2(y) - 1] - y = 1$. Assuming $N(p)$ and $[S^2(y) - p] - y = 1$, $[S^2(y) - S(p)] - y = [[S^2(y) - p] - y] - 1$ (11.3), $= 1 - 1$ (by the hyp.), conv 1. Hence, by induction, $N(p) \supset_p \cdot [S^2(y) - p] - y = 1$. This with $N(S^2(y) - x)$ yields $[S^2(y) - \cdot S^2(y) - x] - y = 1$, and, by 12.6, $x - y = 1$.

12.17: $$[x < S(y) \cdot y < S(x)] \supset_{xy} \cdot x = y.$$

Proof. Assume $x < S(y) \cdot y < S(x)$. $x < S(y)$, 12.11 $\vdash S(x) < S^2(y)$. (1) $2 + y$ conv $S^2(y)$, $= [S^2(y) - S(x)] + S(x)$ (12.5, $S(x) < S^2(y)$), $= [S^2(y) - S(x)] + [S(x) - y] + y$ (12.5, $y < S(x)$). Hence, by 11.4, (2) $2 = [S^2(y) - S(x)] + \cdot S(x) - y$. (3) $S^2(y) - S(x) = [[S^2(y) - S(x)] + \cdot S(x) - y] - \cdot S(x) - y$ (11.2), $= 2 - \cdot S(x) - y$ (by (2)), $= 2 - 1$ (12.16, $S(x) < S^2(y)$), conv 1. (4) $2 + x$ conv $1 + S(x)$, $= [S^2(y) - S(x)] + S(x)$ (by (3)), $= 2 + y$ (as in (1)). Hence, by 11.4, $x = y$.

12.18: $$[x > 1] \supset_x \cdot N(y) \supset_y \cdot x - y < x.$$

---

* Or let $\leq \to \lambda xy \cdot x < S(y)$ and abbreviate $\{\leq\}(x, y)$ to $x \leq y$. Similarly below.

Proof. Assume $x > 1$. (1) $1 - 1$ conv $1, < S(1)$. Assuming $N(p)$, $S(p) - 1 = p, < S^2(p)$. Hence, by induction, $N(p) \supset_p \cdot p - 1 < S(p)$. (2) $x - 1 < S(x - 1), = x$ (12.5). (3) Assuming $N(y)$ and $x - y < x$, $x - S(y) = [x - y] - 1$ (11.3), $\leq x - y$ (by means of (1)), $< x$ (hyp. induction). From (2) and (3), by induction, $N(y) \supset_y \cdot x - y < x$.

12.19: $N(k) \supset_k \cdot [\rho < S(k) \supset_\rho \phi(\rho)]\,[\rho > k \supset_\rho \phi(\rho)] \supset_\phi \cdot N(x) \supset_x \phi(x)$.

Proof. Let $\mathfrak{D}_\phi(a) \rightarrow [\rho < S(a) \supset_\rho \phi(\rho)][\rho > a \supset_\rho \phi(\rho)]$. (1) $\vdash \Sigma\phi \cdot \mathfrak{D}_\phi(1)$. By induction and Theorem I, $\mathfrak{D}_\phi$ (1) $\supset_\phi \cdot N(x) \supset_x \phi(x)$. (2) Assume $N(k)$ and $\mathfrak{D}_\phi(k) \supset_\phi \cdot N(x) \supset_x \phi(x)$. Then $\Sigma\phi \cdot \mathfrak{D}_\phi(S(k))$. Assume $\mathfrak{D}_\phi(S(k))$. (a) $\rho < S(k) \vdash \rho < S^2(k)$; and hence, by the first factor of $\mathfrak{D}_\phi(S(k))$, and Theorem I, $\rho < S(k) \supset_\rho \phi(\rho)$. (b) $N(k) \vdash 1 + k < S^2(k)$, whence, by the first factor of $\mathfrak{D}_\phi(S(k))$, $\phi(1 + k)$. Assuming $N(p)$, we have $S(p) + k > S(k)$, whence, by the second factor of $\mathfrak{D}_\phi(S(k))$, $\phi(S(p) + k)$. By induction, $N(p) \supset_p \phi(p + k)$. Thence, assuming $\rho > k$, we obtain $\phi([\rho - k] + k)$, and, by 12.5, $\phi(\rho)$. By Theorem I, $\rho > k \supset_\rho \phi(\rho)$. (c) $\mathfrak{D}_\phi(k) \supset_\phi \cdot N(x) \supset_x \phi(x)$ with (a) and (b) yields $N(x) \supset_x \phi(x)$. By Theorem I, $\mathfrak{D}_\phi(S(k)) \supset_\phi \cdot N(x) \supset_x \phi(x)$. (3) 12.19 follows from (1) and (2) by induction.

13. The lesser and greater of two positive integers. Let $\min \rightarrow \lambda xy \cdot S(y) - \cdot S(y) - x$ and $\max \rightarrow \lambda xy \cdot [x + y] - \min(x, y)$.

13.1: $N(x)N(y) \supset_{xy} N(\min(x, y))$.
13.2: $[N(y) \cdot x < S(y)] \supset_{xy} \cdot \min(x, y) = x$.
13.3: $N(x) \supset_x \cdot N(y) \supset_y \cdot \min(x, y) = \min(y, x)$.
13.4: $N(x)N(y) \supset_{xy} \cdot \min(x, y) < S(y)$.
13.5: $N(x)N(y)N(z) \supset_{xyz} \cdot \min(z + x, z + y) = z + \min(x, y)$.

Proofs. 3.2, 11.1 $\vdash$ 13.1. 12.6 $\vdash$ 13.2. 13.3 may be established by an application of 12.19, since, assuming $N(x)$, (1) assuming $y < S(x)$, $\min(x, y)$ conv $S(y) - \cdot S(y) - x, = S(y) - 1$ (by 12.16, since $y < S(x)$, 12.11 $\vdash S(y) < S^2(x)$), $= y, = \min(y, x)$ (13.2), and (2) assuming $y > x$, a like series of steps takes us from $\min(y, x)$ to $\min(x, y)$. 12.18, 12.8 $\vdash$ 13.4. $\vdash$ 13.5, since, assuming $N(x)N(y)N(z)$, $\min(z + x, z + y)$ conv $S(z + y) - \cdot S(z + y) - \cdot z + x, = [z + S(y)] - \cdot S(y) - x$ (11.2, 11.3), $= z + \cdot S(y) - \cdot S(y) - x$ (12.7, 12.18, 12.8), conv $z + \min(x, y)$.

14. Proof by cases. We now establish theorems which connect the present theory with that of §§ C7, C9, C10. $M \rightarrow \lambda\mu \cdot \phi(1)\phi(2) \supset_\phi \phi(\mu)$,

and $[\boldsymbol{P}] \equiv_{x_1 \ldots x_n} \cdot \boldsymbol{Q}$ will be used as an abbreviation for $[[\boldsymbol{P}] \supset_{x_1 \ldots x_n} \cdot \boldsymbol{Q}]$ $[[\boldsymbol{Q}] \supset_{x_1 \ldots x_n} \cdot \boldsymbol{P}]$.*

14.1: $\quad M(x) \equiv_x \cdot x < 3.$

Proof. C7.1 $\vdash \Sigma x \cdot M(x)$. $1 < 3$, $2 < 3$, $M(x) \vdash x < 3$, by Rule V. Hence $M(x) \supset_x \cdot x < 3$. Conversely, assume $x < 3$ and $\phi(1)\phi(2)$. Then $\phi(\min(1, 2))$ ($\phi(1)$, 13.2); and assuming $N(p)$, $\phi(\min(S(p), 2))$ ($\phi(2)$, 13.2, 13.3, 12.4). Hence, by induction, $N(p) \supset_p \phi(\min(p, 2))$. Thence, since $x < 3 \vdash N(x)$, we obtain $\phi(\min(x, 2))$, and, by 13.2, $\phi(x)$.

Note that $M(\boldsymbol{x}) \vdash N(\boldsymbol{x})$ (by 14.1, 15).

Let $[\boldsymbol{x}] \circ [\boldsymbol{y}] \to \min(\boldsymbol{x}, \boldsymbol{y})$. $\circ$ multiplies 1's and 2's as 0's and 1's resp.

14.2: $\quad M(x)M(y) \supset_{xy} M(x \circ y).$

14.3: $\quad M(x)M(y) \supset_{xy} \cdot x \circ y = y \circ x.$

14.4: $\quad M(y) \supset_y \cdot 1 \circ y = 1.$

14.5: $\quad M(y) \supset_y \cdot 2 \circ y = y.$

14.6: $\quad [M(x)M(y) \cdot x \circ y = 2] \equiv_{xy} \cdot [x = 2]\,[y = 2].$

Proofs. 14.5 follows from $2 \circ 1 = 1$ and $2 \circ 2 = 2$ by the definition of M. For 14.6a (i.e. the first factor of 14.6), assume $M(x)M(y) \cdot x \circ y = 2$. Then, by 13.4, $x \circ y < S(y)$. Hence $2 < S(y)$. Also, by 14.1, $y < S(2)$. Hence, by 12.17, $y = 2$. Likewise $x = 2$ (cf. 14.3).

Let $\epsilon \to \lambda xy \cdot \min(2, S(x) - y)$, and abbreviate $\epsilon(\boldsymbol{x}, \boldsymbol{y})$ to $\epsilon_{\boldsymbol{y}}^{\boldsymbol{x}}$.

14.7: $\quad N(x)N(y) \supset_{xy} M(\epsilon_y^x).$

14.8: $\quad [N(y) \cdot x < S(y)] \equiv_{xy} \cdot N(x)N(y) \cdot \epsilon_y^x = 1.$

14.9: $\quad x > y \equiv_{xy} \cdot N(x)N(y) \cdot \epsilon_y^x = 2.$

Proofs. 13.3, 13.4, 14.1 $\vdash$ 14.7. 12.16, 12.11 $\vdash$ 14.8a; and for 14.9a we have, assuming $x > y$, $\epsilon_y^x = \min(2, S(x) - y)$, $= \min(2, S(x - y))$ (12.7), $= 2$ (by 13.2, since 12.4 $\vdash S^2(x - y) > 2$). 14.8b and 14.9b we prove as follows: By C7I, there exists a formula $\mathfrak{B}$ such that $\mathfrak{B}(1)$ conv $\lambda ab \cdot a < S(b)$ and $\mathfrak{B}(2)$ conv $\lambda ab \cdot a > b$. Assuming $N(y)$, then (1) assuming $x < S(y)$ we infer $\epsilon_y^x = 1$ by 14.8a, and hence $\mathfrak{B}(\epsilon_y^x, x, y)$, and (2) assuming $x > y$, we infer $\epsilon_y^x = 2$ by 14.9a, and hence $\mathfrak{B}(\epsilon_y^x, x, y)$. By 12.19 and Theorem I, $N(y) \supset_y \cdot N(x) \supset_x \mathfrak{B}(\epsilon_y^x, x, y)$. This lemma enables us to infer $x < S(y)$ from $N(x)N(y) \cdot \epsilon_y^x = 1$ and $x > y$ from $N(x)N(y) \cdot \epsilon_y^x = 2$.

* Cf. Church, 1932, p. 355.

Let $\delta \to \lambda xy \cdot 4 -\cdot \epsilon_y^x + \epsilon_x^y$, and abbreviate $\delta(x, y)$ to δ_y^x.

14.10: $N(x)N(y) \supset_{xy} M(\delta_y^x)$.

14.11: $N(x)N(y) \supset_{xy} \cdot \delta_y^x = \delta_x^y$.

14.12: $x < y \supset_{xy} \cdot \delta_y^x = 1$.

14.13: $[N(y) \cdot x < S(y) \cdot \delta_y^x = 1] \supset_{xy} \cdot x < y$.

14.14: $[N(x)N(y) \cdot x = y] \equiv_{xy} \cdot N(x)N(y) \cdot \delta_y^x = 2$.

Proofs. Assuming $N(x)N(y)$, $\delta_y^x = 4 -\cdot \epsilon_y^x + \epsilon_x^y$, $= [4 - \epsilon_y^x] - \epsilon_x^y$ (11.3), $< 4 - \epsilon_y^x$ (by 12.18, since $4 - \epsilon_y^x = 1 + \cdot 3 - \epsilon_y^x$ (12.7, 14.1, 14.7), > 1 (12.10)), ≤ 3 (12.18), and hence, by 14.1, $M(\delta_y^x)$. 5.4 $\vdash$ 14.11. Assuming $x < y$, $\delta_y^x = 4 -\cdot \epsilon_y^x + 2$ (14.9), $= [4 - 2] - \epsilon_y^x$ (11.3), conv $2 - \epsilon_y^x$, $= 1$ (12.16). Assuming $N(y) \cdot x < S(y) \cdot \delta_y^x = 1$, $\epsilon_x^y = 3 -\cdot 3 - \epsilon_x^y$ (12.6, 14.1, 14.7), $= 3 -\cdot 4 -\cdot 1 + \epsilon_x^y$ (11.2, 11.3), $= 3 -\cdot 4 -\cdot \epsilon_y^x + \epsilon_x^y$ (14.8), $= 3 - \delta_y^x$ (def.), $= 3 - 1$ (hyp.), conv 2; and hence, by 14.9, $x < y$. Assuming $N(x)N(y) \cdot x = y$, then $x < S(y)$ and, by 14.8, $\epsilon_y^x = 1$; also $\epsilon_x^y = 1$; hence $\delta_y^x = 4 -\cdot 1 + 1$, conv 2. It remains to establish 14.14b, which we do as follows. Assume $N(x)N(y) \cdot \delta_y^x = 2$. By 14.7, $M(\epsilon_y^x)$. (1) Assuming $\epsilon_y^x = 1$, then $x < S(y)$ (14.8). (2) Assuming $\epsilon_y^x = 2$, then $x > y$ (14.9); hence $\delta_y^x = 1$ (14.11, 14.12). Hence, by C10II, $x < S(y)$. Similarly, $y < S(x)$. Hence, by 12.17, $x = y$.

Henceforth we use tacitly results of this section in conjunction with C9I and C10II. The headings "Case 1," "Subcase a," etc., will indicate applications of C9I or C10II.

PRINCETON UNIVERSITY.

ON THE COOLING OF THE EARTH.

By Arnold N. Lowan.

Abstract.

In an earlier paper * the thermal history of the earth was investigated under the assumption that the earth is a radioactive sphere, the surface of which is permanently at 0°. The present paper introduces the further refinement of taking into account the heat evolved as a result of the contraction accompanying the cooling of the earth. The method is essentially that based on the use of the Laplace transformation as presented in the earlier paper above mentioned in conjunction with the theory of biorthogonal sets of functions and the corresponding integral equations with asymmetric kernels.

The theory underlying the subsequent derivations is due to the work of J. M. C. Duhamel † who nearly a hundred years ago has set up the differential equations governing the cooling of a solid when account is taken of the heat set free as a result of its accompanying contraction. As an illustration of his theory, Duhamel has obtained the complete solution for the problem of the cooling of a sphere (with due account of contraction) the surface of which is permanently at 0°. The present paper contemplates the generalization of Duhamel's problem to the case where the sphere is radioactive.

As in A. N. L. it will be assumed that the sphere (earth) is centrally symmetrical with regard to the initial temperature distribution, the distribution of radioactive matter and that the physical constants do not vary with depth.

In view of the results in J. D. and A. N. L., it is clear that the mathematical formulation of our problem is as follows:

$$(1) \qquad \frac{\partial}{\partial t}T(r,t) = k\left(\frac{\partial^2}{\partial r^2}T(r,t) + \frac{2}{r}\frac{\partial}{\partial r}T(r,t)\right) + \frac{C_p - C_v}{3C_v\delta}\frac{\partial}{\partial t}\left(3\psi + r\frac{\partial \psi}{\partial r}\right) + \phi(r,t)$$

$$(2) \qquad r\frac{\partial^2}{\partial r^2}\psi(r,t) + 4\frac{\partial}{\partial r}\psi(r,t) - \frac{5\delta}{3}\frac{\partial}{\partial r}T(r,t) = 0$$

* Arnold N. Lowan, "On the cooling of a radioactive sphere," *Physical Review*, November 1, 1933. This paper will be referred to as A. N. L.

† *Journal de l'école Polytechnique* 1837. This paper will be referred to as J. D.

$$\lim_{t\to 0} T(r,t) = f(r) \tag{3}$$

$$\lim_{t\to 0} \psi(r,t) = 0 \tag{4}$$

$$T(R,t) = 0 \tag{5}$$

$$5\psi(r,t) + 3r(\partial/\partial r)\psi(r,t) = 0 \qquad \text{for } r = R \tag{6}$$

where the significance of the symbols employed is as follows:

$T(r,t)$ = temperature at time t, and distance r from center;

R = radius of the earth;

k = thermal diffusivity = ratio between the thermal conductivity K and the product of the density ρ and the specific heat C_v (under constant volume);

C_p = specific heat under constant pressure;

δ = coefficient of linear thermal expansion;

$\phi(r,t) = 1/\rho C_v \times$ heat generated per unit time per unit volume by the radioactive matter;

$r\psi$ = increment of the radius vector which initially had the value r.

It may be remarked that if $C_p = C_v$, the problem becomes identical with that treated in A. N. L.; further the system (1) to (6) is identical with that treated by Duhamel, except for the term $\phi(r,t)$ in the differential equation (1). Thus the solution to be derived must reduce to that given in A. N. L. if $C_p = C_v$ and to that given by Duhamel if $\phi(r,t) = 0$

Consider the differential equation

$$r\,\partial^2\psi/\partial^2 + 4\,\partial\psi/\partial r = 0. \tag{7}$$

If we assume a solution of (7) in the form $\psi = Ar^m$ it is readily found that $m(m+3) = 0$. Thus the general solution of (7) is in the form $\psi = A + Br^{-3}$. Starting with this solution of (7) the solution of (2) satisfying the boundary condition (6) and such that $r\psi$ is finite for $r = 0$ is readily found by the method of variation of parameters in the form:

$$\psi(r) = \frac{5\delta}{r^3}\int_0^r \rho^2 T(\rho)\,d\rho + \frac{4\delta}{3R^3}\int_0^R \rho^2 T(\rho)\,d\rho \tag{8}$$

whence ultimately

$$\frac{\partial}{\partial t}\left(3\psi + r\frac{\partial\psi}{\partial r}\right) = \frac{5\delta}{3}\frac{\partial T}{\partial t} + \frac{4\delta}{3}\int_0^R \rho^2 \frac{\partial T}{\partial t}\,d\rho. \tag{9}$$

In view of (8), equation (1) becomes

$$\frac{\partial T}{\partial t} = k\left(\frac{\partial^2 T}{\partial r^2} + \frac{2}{r}\frac{\partial T}{\partial r}\right) + \frac{5}{9}\frac{C_v - C_p}{C_v}\frac{\partial T}{\partial t} + \frac{4}{3}\frac{C_v - C_p}{R^3 C_v}\int_0^R \rho^2 T(\rho)\,d\rho + \phi(r,t). \tag{10}$$

If we make the substitution:

$$T(r,t) = (1/r)u(r,t) + rf(r). \tag{11}$$

It is ultimately found that the function $u(r,t)$ must satisfy the system of equations.

$$A^2 \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial r^2} + F''(r) - B^2 r \int_0^R \rho \frac{\partial u}{\partial t}\, d\rho + \frac{r}{k}\phi(r,t) \tag{12}$$

$$\lim_{t\to 0} u(r,t) = 0 \tag{13}$$

$$u(0,t) = u(R,t) = 0 \tag{14-15}$$

where we have put:

$$A^2 = \frac{1}{k}\left(1 + \frac{5}{9}\frac{C_p - C_v}{C_v}\right);\quad B^2 = \frac{4}{3}\frac{C_p - C_v}{R^3 k C_v};\quad F(r) = rf(r). \tag{16}$$

In accordance with the method employed in A. N. L. we now operate on (12), (14) and (15) by the "Laplace operator" defined as follows:

$$L\{u(r,t)\} = \int_0^\infty e^{-\lambda t} u(r,t)\, dt = y(r,\lambda) \tag{17}$$

$$L\{\phi(r,t)\} = \psi(r,\lambda). \tag{18}$$

If we make use of the identity *

$$L\{(\partial/\partial t)u(r,t)\} = \lambda L\{u(r,t)\} - u(r,0) = \lambda y - F(r) \tag{19}$$

it is ultimately found that the function $y(r,\lambda)$ must satisfy the equations:

$$y'' + \alpha^2 y = -\alpha^2 \frac{B^2}{A^2} r \int_0^R \rho y(\rho)\, d\rho - \sigma(r,\lambda) \tag{20}$$

$$y(0) = y(R) = 0 \tag{21-22}$$

where we have put:

$$-\lambda A^2 = \alpha^2 \qquad \sigma(r,\lambda) = (1/\lambda)F''(r) + (r/k)\psi(r,\lambda). \tag{23}$$

To solve the system (20) to (22) we start with the simpler system

$$y'' + \alpha^2 y = -\alpha^2 \frac{B^2}{A^2} r \int_0^R \rho y(\rho)\, d\rho = -\alpha^2 C r \quad \text{(say)} \tag{24}$$

$$y(0) = y(R) = 0. \tag{20-21}$$

If we consider provisionally C as a given constant, the solution of (24), (20) and (21) is readily found in the form:

* See A. N. L.

$$y(r,\alpha) = C\left(-r + \frac{R}{\sin R\alpha}\sin r\alpha\right) \tag{25}$$

whence

$$\int_0^R \rho y(\rho)\,d\rho = -\frac{CR^3}{3} + \frac{CR}{\sin R\alpha}\cdot\frac{\sin R\alpha - R\alpha\cos R\alpha}{\alpha^2}. \tag{26}$$

By the definition of the constant C from (24) we have:

$$C = \frac{B^2}{A^2}\int_0^R \rho y(\rho)\,d\rho.$$

Therefore in view of (26) we ultimately get:

$$1 - R\alpha\cot R\alpha = R^2\alpha^2\left(\frac{1}{3} + \frac{A^2}{B^2R^3}\right). \tag{27}$$

Thus the solution of the system (24), (20) and (21) is given by (25) where the α's are the roots of the transcendental equation (27) and where now C is an arbitrary constant.

Consider now the system

$$z'' + \beta^2 z = 0 \tag{28}$$

$$z(0) = 0 \tag{29}$$

where the solution of (28) will be subjected to the condition of satisfying an additional boundary condition to be subsequently determined.

Let α_m be a characteristic value of the system S consisting of (24), (20) and (21) and $y_m(r)$ the corresponding characteristic function. Let similarly β_n designate a characteristic value of the system T consisting of (28), (29) and the as yet undetermined additional boundary condition and $z_n(r)$ the corresponding characteristic function. Then by the familiar "Green process" and in view of the boundary conditions (20), (21) and (29) we get

$$y'_m(R)z_n(R) + (\alpha_m^2 - \beta_n^2)\int_0^R y_m(\rho)z_n(\rho)\,d\rho + \alpha_m^2 C_m\int_0^R \rho z_n(\rho)\,d\rho = 0 \tag{30}$$

where C_m is the value of the constant C as defined in (24) for $y = y_m$.

It is therefore clear that if we put the condition:

$$y'_m(R)z_n(R) + \alpha_m^2 C_m\int_0^R \rho z_n(\rho)\,d\rho = 0 \tag{31}$$

the functions $y_m(r)$ and $z_n(r)$ will satisfy the "biorthogonal" condition

$$\int_0^R z_n(\rho)y_m(\rho)\,d\rho = 0 \qquad \text{if } m \neq n. \tag{32}$$

It will be shown that (31) is identically satisfied provided that the α_n's and β_n's satisfy the transcendental equation (27).

We evidently have

$$\int_0^R \rho z_n(\rho)\,d\rho = -\frac{1}{\beta_n^2}\int_0^R \rho z_n''(\rho)\,d\rho \tag{33}$$

$$= \frac{1}{\beta_n^2}\{z_n(R) - z_n(0) - Rz'_n(R)\}.$$

The solution of (28) satisfying (29) is

$$z_n(r) = M_n \sin \beta_n r \tag{34}$$

where M_n is an arbitrary constant. Further from (25) we get:

$$y'_m(R) = C_m(-1 + R\alpha_m \cot R\alpha_m). \tag{35}$$

With the aid of (33), (34) and (35) the boundary condition (31) becomes

$$\left\{\frac{1}{\alpha_m^2}(-1 + R\alpha_m \cot R\alpha_m) - \frac{1}{\beta_n^2}(-1 + R\beta_n \cot R\beta_n)\right\} \sin R\beta_n = 0. \tag{36}$$

In view of (27) it is clear that (36) and therefore (31) are identically satisfied provided the α_n's and β_n's are the roots of the transcendental equation (27).

From (35) and (27) we get:

$$\frac{y'_m(R)}{\alpha_m^2 C_m} = -R^2\left(\frac{1}{3} + \frac{A^2}{B^2R^3}\right) = -h \quad \text{(say)}. \tag{37}$$

Thus (31) may be written in the more convenient form:

$$hz_n(R) = \int_0^R \rho z_n(\rho)\,d\rho. \tag{31'}$$

This is the additional boundary condition for the function $z_n(r)$, previously referred to.

Summarizing the results found thus far the solution of the systems S and T are:

$$y_n(r, \alpha) = M_n\left(-r + \frac{R}{\sin R\alpha_n}\sin r\alpha_n\right) \tag{25'}$$

$$z_n(r, \alpha) = N_n \sin r\alpha_n \tag{34'}$$

where M_n and N_n are arbitrary constants and α_n is a root of the transcendental equation (27).

In the subsequent derivations it will be assumed that the solutions (25') and (34') have been normalized in accordance with:

$$\int_0^R y_n(r) z_n(r)\, dr = 1 \tag{38}$$

whence

$$M_n N_n = \frac{1}{\int_0^R \sin \alpha_n \rho \left(-\rho + \frac{R}{\sin R\alpha_n} \sin \rho\alpha_n\right) d\rho}. \tag{39}$$

We now return to the system S and define the "Green function" $K(r, \xi)$ as the solution of the system satisfying the discontinuity condition:

$$\operatorname*{Lim}_{\epsilon \to 0} \frac{\partial}{\partial r} K(r, \xi) \Big|_{\xi+\epsilon}^{\xi-\epsilon} = -1. \tag{40}$$

With the aid of (40) it can be readily verified that the expression:

$$y(r, \alpha) = \int_0^R K(r, \xi) \sigma(\xi, \lambda)\, d\xi \tag{41}$$

is the complete solution of the differential equation (20) satisfying the boundary conditions (21) and (22).

Consider now the differential equation satisfied by $K(r, \xi)$

$$K''(r, \xi) + \alpha^2 K(r, \xi) = -\alpha^2 \frac{B^2}{A^2} r \int_0^R \rho K(\rho, \xi)\, d\rho \tag{42}$$

in conjunction with the differential equation:

$$z_n'' + \beta_n^2 z_n = 0 \tag{28'}$$

satisfied by $z_n(r)$. If we perform the "Green process" on (42) and (28') and make use of the discontinuity condition (40), we ultimately get:

$$(\alpha^2 - \alpha_n^2) \int_0^R K(\rho, \xi) z(\rho)\, d\rho + z_n(\xi) = 0 \tag{43}$$

which is an integral equation satisfied by the characteristic functions z_n.

In entirely similar fashion we define the Green function $G(r, \xi)$ as the solution of the system T satisfying the discontinuity condition

$$\operatorname*{Lim}_{\epsilon \to 0} \frac{\partial}{\partial r} G(r, \xi) \Big|_{\xi+\epsilon}^{\xi-\epsilon} = -1 \tag{44}$$

if then we perform the Green process on (42) and:

$$G'' + \alpha^2 G = 0 \tag{45}$$

and take into account the boundary and discontinuity conditions satisfied by the Green functions $K(r, \xi)$ and $G(r, \xi)$ we readily find the important relation:

$$K(r,\xi) = G(\xi,r) \tag{46}$$

Furthermore by a method entirely similar to that which has led to the integral equation (43) satisfied by the characteristic functions z_n, we obtain an integral equation satisfied by the characteristic functions y in the form:

$$(\alpha^2 - \alpha_n{}^2)\int_0^R G(\rho,\xi)y(\rho)\,d\rho + y_n(\xi) = 0. \tag{47}$$

Thus the functions y_n and z_n are solutions of the integral equations (43) and (47). Since they satisfy the "biorthogonal" condition (32) they are said to form a biorthogonal set. As previously stated these functions are assumed to be normalized in accordance with (38).

The fact that the functions y_n and z_n satisfy the integral equations (43) and (47) with asymmetric kernels (in view of (46)) is of paramount importance for our purpose. Indeed it is well known * that for these asymmetric kernels the bilinear expansion formulae

$$K(r,\xi) = \sum_{n=1}^{\infty} \frac{y_n(r)z_n(\xi)}{\alpha_n{}^2 - \alpha^2} \tag{48}$$

$$G(r,\xi) = \sum_{n=1}^{\infty} \frac{y_n(\xi)z_n(r)}{\alpha_n{}^2 - \alpha^2} \tag{49}$$

are valid, where the second members are uniformly convergent in the interval $O - R$ under consideration.

In view of (48) and of the significance of the function $\sigma(r,\lambda)$ from (23) our solution (41) becomes:

$$\begin{aligned} y(r,\lambda) = &-\sum_{n=1}^{\infty} \frac{y_n(r)}{\alpha_n{}^2 - \alpha^2}\,\frac{A^2}{\alpha^2}\int_0^R F''(\xi)z_n(\xi)\,d\xi \\ &+ \sum_{n=1}^{\infty} \frac{y_n(r)}{\alpha_n{}^2 - \alpha^2}\int_0^R \xi\psi(\xi,\lambda)z_n(\xi)\,d\xi. \end{aligned} \tag{50}$$

If the integral in the first term of (50) is twice integrated by parts we get:

$$\int_0^R F''(\xi)z_n(\xi)\,d\xi = \{z_n(R)F'(R) - z'_n(R)F(R)\} + \int_0^R z_n''(\rho)F(\rho)\,d\rho. \tag{51}$$

Consider the contribution to $y(r,\lambda)$ as given in (50) arising from the expression in brackets in (51). With the aid of (48) it is clear that this contribution may be written in the form:

$$\frac{A^2}{\alpha^2}F'(R)\left\{K(r,\xi)\right\}_{\xi=R} - \frac{A^2}{\alpha^2}F(R)\left\{\frac{\partial}{\partial\xi}K(r,\xi)\right\}_{\xi=R} \tag{52}$$

* A. Kneser, *Integralgleichungen*, Section 6.

If we designate by y_1 and y_2 the analytical expressions of the Green function $K(r,\xi)$ in the intervals $(O-\xi)$ and $(\xi-R)$ we may write:

$$(53)\qquad \begin{aligned} y_1 &= -C_1 r + M_1 \sin \alpha r + N_1 \cos \alpha r \\ y_2 &= -C_2 r + M_2 \sin \alpha r + N_2 \cos \alpha r. \end{aligned}$$

The six coefficients in y_1 and y_2 may be computed from the following conditions which y_1 and y_2 must evidently satisfy

$$\begin{aligned} &y_1(0) = 0; \quad y_2(R) = 0; \quad y_1(\xi) = y_2(\xi) \\ &y'_1(r,\xi) - y'_2(r,\xi) = 1 \qquad \text{for } r = \xi. \end{aligned}$$

$$(54)\qquad \int_0^R \rho y_1(\rho,\xi)\,d\rho = \frac{A^2}{B^2}\,C_1; \quad \int_0^R \rho y_2(\rho,\xi)\,d\xi = \frac{A^2}{B^2}\,C_2.$$

It may be easily verified that for $\xi = R$ the above system of equations yields $C_1 = M_1 = N_1 = 0$. Accordingly the first term in (52) vanishes identically. The vanishing of the second term may be proven in a similar manner. Thus (51) reduces to its second term and therefore

$$(51')\qquad \int_0^R F''(\xi) z_n(\xi)\,d\xi = \int_0^R z_n''(\xi) F(\xi)\,d\xi = -\alpha_n^2 \int_0^R F(\xi) z_n(\xi)\,d\xi.$$

With the aid of (51') and (23) our solution (50) becomes after an obvious transformation:

$$(55)\qquad \begin{aligned} y(r,\lambda) = &-\frac{1}{\lambda}\sum_{n=1}^{\infty} y_n(r) \int_0^R F(\rho) z_n(\rho)\,d\rho + \sum_{n=1}^{\infty} \frac{y_n(r)}{(\alpha_n/A)^2 + \lambda} \int_0^R F(\rho) z_n(\rho)\,d\rho \\ &+ \frac{1}{kA^2} \sum_{n=1}^{\infty} \frac{y_n(r)}{(\alpha_n/A)^2 + \lambda} \int_0^R \xi \psi(\xi,\lambda) z_n(\xi)\,d\xi. \end{aligned}$$

The complete solution of our problem is now obtained by subjecting (55) to the inverse Laplace transformation L^{-1}. Since we evidently have $L^{-1}\{\lambda\} = 1$ the contribution to $T(r,t)$ arising from the first term in (55) reduces to $-F(r)$, provided we assume that the function $F(r)$ is piecewise continuous and twice differentiable, in which case we have:

$$F(r) = \sum_{n=1}^{\infty} y_n(r) \int_0^R F(\rho) z_n(\rho)\,d\rho.$$

Further since $L^{-1}\left\{\frac{1}{(\alpha_n/A)^2 + \lambda}\right\} = e^{-(\alpha_n/A)^2 t}$ and $L^{-1}\{\psi(r,\lambda)\} = \phi(r,t)$ we have by virtue of Borel's theorem *

* See A. N. L.

$$(56) \qquad L^{-1}\left\{\frac{1}{(\alpha_n/A)^2+\lambda}\psi(r,\lambda)\right\} = e^{-(\alpha_n/A)^2 t}\int_0^t \phi(r,\eta)\, e^{(\alpha_n/A)^2\eta}\, d\eta.$$

The inversion of (55) thus finally yields:

$$T(r,t) = \frac{1}{r}\sum_{n=1}^{\infty} e^{-(\alpha_n/A)^2 t}\, y_n(r)\Big[\int_0^R F(\rho) z_n(\rho)\, d\rho + \frac{1}{kA^2}\int_0^R \rho z_n(\rho)\left\{\int_0^t \phi(\rho,\eta)\, e^{(\alpha_n/A)^2\eta}\, d\eta\right\} d\rho\Big]$$

substituting for $y_n(r)$ and $z_n(r)$ the values from (25′) and (34′) and computing the value of the normalizing factor from (39) the last equation becomes after a slight transformation:

$$T(r,t) = \frac{2}{r}\sum_{n=1}^{\infty} \frac{e^{-(\alpha_n/A)^2 t}\left(\sin \alpha_n r - \frac{r}{R}\sin R\alpha_n\right)}{R - 2\,\frac{\sin^2 R\alpha_n}{R{\alpha_n}^2} + \frac{1}{2\alpha_n}\sin 2R\alpha_n}\int_0^R \rho f(\rho)\sin \rho\alpha_n d\rho$$

$$+ \frac{2}{A^2 k r}\sum_{n=1}^{\infty} \frac{e^{-(\alpha_n/A)^2 t}\left(\sin r\alpha_n - \frac{r}{R}\sin R\alpha_n\right)}{R - 2\,\frac{\sin^2 R\alpha_n}{R{\alpha_n}^2} + \frac{1}{2\alpha_n}\sin 2R\alpha_n}$$

$$\times \int_0^R \rho \sin \rho\alpha_n \left\{\int_0^t \phi(\rho,\eta)\, e^{(\alpha_n/A)^2\eta}\, d\eta\right\} d\rho.$$

In (57) and (8) we have the complete solution of our problem.

It is readily ascertained that for $C_p = C_v$ (57) reduces to the solution given in A. N. L. Furthermore for $\phi(r,t) = 0$, (57) yields at once the solution given by Duhamel.

INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.